TALLINN UNIVERSITY OF TECHNOLOGY
School of Business and Governance
Department of Law

Aleksandr Kesa, Tanel Kerikmäe

ARTIFICIAL INTELLIGENCE AND THE GDPR: INEVITABLE NEMESES?

Master's thesis
MA Law and Technology

Supervisor: Tanel Kerikmäe, PhD

Tallinn 2020

TABLE OF CONTENTS

ABSTRACT

The rapid development of computer technology over the past decades has brought about countless benefits across industries and social benefits as well – constant interpersonal connectivity is facilitated through numerous communication channels and social media outlets, energy-producing enterprises employ complex machinery management systems for increased efficiency, ease of access and safety, hedge funds make use of HFT algorithms to engage in trades happening at the fraction of a second, while medical professionals use predictive technologies to diagnose diseases and forecast viral outbreaks. Widespread adoption of technology necessitated the creation of regulatory frameworks that would ensure the safeguarding of rights and regulatory and judicial supervision over the exploitation of high technology. One such framework is the GDPR, created due to the need for a comprehensive, contemporary legal regime governing the processing of personal data in a time when such data has become a commodity that is traded and sold in return for services or financial gain. However, in the author's view, the GDPR suffers in terms of efficacy in the context of artificial intelligence-based technologies, and full compliance of data controllers and processors employing such technologies is unlikely to be achieved, particularly in regards to the right to information, the general principle of transparency and the right to erasure. The paper provides an overview of these issues, including a discussion on the movement towards a regime of data ownership, and proposes legislative amendments as an effective method of mitigating these drawbacks.

Keywords: GDPR, AI, machine learning, right to erasure, personal data processing, compliance

INTRODUCTION

The increasingly rapid pace at which personal computing technology has developed over the course of the previous several decades, beginning with the 1977 release of one of the first successfully mass-marketed personal computers, and the various technical innovations in the field, have made technology accessible and affordable to private individuals, unaffiliated with large research centers or enterprises, which enabled an increasingly wide range of persons to be part of and contribute to the growth of personal computing, which, in turn, only accelerated the advancement of technology. Developments in hardware engineering and software practice made computing systems capable of executing more broad and demanding programs and tasks – as processing and graphical capabilities grew, computers shifted from being used solely for calculations and primitive text input, to being used for engineering, architectural design, medicine, and, especially relatively recently, to big data processing and development of artificial intelligence systems. This paper will focus on researching this last application – artificial intelligence (or AI) – in particular under the lens of European Union data protection regulations, and with the goal of determining whether the principles and provisions of such regulations, in particular the General Data Protection Regulation (hereinafter GDPR), are compatible with the technical nature of AI. This focus is justified not only by the increasingly common adoption of business solutions based on AI and the all-engulfing scope of the GDPR, but also by the negative attitudes exhibited by some scholars on the compatibility of these two concepts, for example, opinions such as Zarsky's that any such compatibility is null – "The GDPR's provisions are – to borrow a key term used throughout EU data protection regulation – incompatible with the data environment that the availability of Big Data generates."[1]

Coined by the American computer scientist John McCarthy in 1955[2], the term "artificial intelligence" is broad and encompasses a wide variety of characteristics, and ranges from describing such mundane services as product recommendations on e-commerce platforms[3] or entertainment picks for users of film and music streaming services[4] to describing the

[1] Zarsky, Tal Z. Incompatible: The GDPR in the Age of Big Data. Seton Hall Law Review. Vol. 47, issue 4. 2017. p. 996
[2] Smith, Chris et al. The History of Artificial Intelligence. Washington University, December 2006
[3] Tsesis, A. (2019). Data subjects' privacy rights: Regulation of personal data retention and erasure. University of Colorado Law Review, 90(2). p. 601
[4] Government Office for Science. Artificial intelligence: opportunities and implications for the future of decision making. 2015.

employment of algorithms to process vast amounts of data in short periods of time to engage in high-frequency trading[5] or to predict natural disasters or pandemics[6] and even identify ailments by analyzing medical imagery[7]. As succinctly explained by the United Kingdom Government Office for Science, AI "generally (…) refers to the analysis of data to model some aspect of the world. Inferences from these models are then used to predict and anticipate possible future events."[8] In the European Union legal system, one definition for AI may be found in the Ethics Guidelines for Trustworthy AI published by the independent High-Level Expert Group on Artificial Intelligence, where it is described as "software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal."[9] When a particular application of AI leads to decisions or conduct that affects private individuals it is likely that large amounts of personal data pertaining to those individuals was processed (and even compared to data of other individuals, as can be the case of a system designed to make decisions relating to the offering of loans to persons based on their spending and saving habits, and on the comparison of those to habits of persons previously approved for loans); due to the capability of these systems to process large quantities of data quickly (much more efficiently and on a much grander scale than a human being could)[10], as well as due to the fact that the results of this processing either contribute to making decisions concerning individuals, or serve as the basis for automated decision-making by the machine itself, it is necessary to determine whether legislative instruments concerning data usage and privacy create an adequate regime for the regulation of such processing. In addition to concerns regarding opaque[11] processing of personal data, government institutions and academics have raised concerns regarding the potential of automated decision-making solutions to give rise to discrimination[12] and the inability of current

[5] *Ibid.*
[6] Pauwels, E.; Denton, S. W. (2018). The internet of bodies: Life and death in the age of AI. California Western Law Review, 55(1). p. 231
[7] Ferretti, A.; Schneider, M.; Blasimme, A. (2018). Machine learning in medicine: Opening the new data protection black box. European Data Protection Law Review (EDPL), 4(3). p. 321
[8] Government Office for Science. Artificial intelligence: opportunities and implications for the future of decision making. 2015.
[9] Independent High-Level Expert Group on Artificial Intelligence. Ethics Guidelines for Trustworthy AI. 2019
[10] Tene, O.; Polonetsky, J. (2013). Judged by the tin man: Individual rights in the age of big data. Journal on Telecommunications & High Technology Law, 11(2). p. 351
[11] Etzioni, A.; Etzioni, O. (2016). Keeping ai legal. Vanderbilt Journal of Entertainment & Technology Law, 19(1). p. 133
[12] Mazur, J. (2018). Right to access information as collective-based approach to the GDPR's right to explanation in European law. Erasmus Law Review, 11(3). p. 178

frameworks for regulating privacy to address "the sheer complexity and numerosity of cases of algorithmic discrimination."[13] It is also necessary to determine the extent to which legislators can aim to exert control, in particular by referring to the technical nature of these AI systems, as regulating certain characteristics may be either impossible due to technical limitations, or counterproductive due to the indispensability of these characteristics in their unadulterated form for the functioning of the systems.

In particular, of interest are those GDPR (General Data Protection Regulation, the European framework legislative instrument regulating entities processing EU resident data[14] which took effect in May 2018[15]) provisions that govern or concern the definition of personal data and processing thereof, principles of transparency and non-discrimination, and, especially, the right to information, the right to erasure, the right to human intervention in cases of automated decision making and profiling[16], and others. The importance of determining the breadth of the term "personal data" becomes apparent when one discusses deep learning systems; deep learning is a subfield of machine learning, which is in turn a subset of artificial intelligence, and governs multi-layered AI models where each layer performs a specific task of input data analysis or manipulation, and compares the conclusions, or labels assigned to the inputted data, to human-generated correct labels to adjust and improve the automated labelling process[17] – if the adjustments made were based on the data of real, identifiable persons, then it is necessary to make clear whether these inferred conclusions should be regarded as part of that identifiable person's personal data. If this is so, it would be problematic to balance this fact with the right to erasure, as in such a case if an individual requests his or her data to be erased, for example upon the termination of a relationship with the data processor, the data processor would need to not only erase the input data itself, but also prevent the machine from exercising behavior learned by analyzing that person's data – however, as these systems process vast amounts of data at rates faster than a human can, and do so many times over, it may be burdensome, or even impossible, to trace exactly what behavior was learned based on one particular person's personal data.

---

[13] Katyal, S. K. (2019). Private accountability in the age of artificial intelligence. UCLA Law Review, 66(1). p. 100
[14] Tschider, C. A. (2018). Regulating the internet of things: Discrimination, privacy, and cybersecurity in the artificial intelligence age. Denver Law Review, 96(1). p. 131
[15] Selbst, A. D.; Barocas, S. (2018). The intuitive appeal of explainable machines. Fordham Law Review, 87(3). p. 1106
[16] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC, O.J. L 119/40, 119/43, 119/46
[17] Macuga, Tim. What is Deep Learning and How Does It Work? Retrieved from: https://resources.rvhub.org/what-is-deep-learning-and-how-does-it-work/. August 23 2019

Furthermore, the workings of complex deep learning systems that are designed for data-intensive applications have been described as a "black box"[18], meaning that due to the complexity of these systems and the amount of data manipulation carried out by them, it can be impossible to find how exactly a particular decision was made by the system, which puts these systems at odds with the principle of transparency and renders data processors incapable of explaining to a data subject exactly how a particular decision was reached by the automated system. Legislative efforts aiming to regulate these technologies must be "future-proof" – "the law (…) should neither hinder the advance of technology, nor require over-frequent revision to tackle such progress."[19]

With the aforementioned taken into account, the research will focus on determining whether the European Union data protection framework is adequate to efficiently regulate the using of artificial intelligence systems that process the personal data of European Union persons – efficiency in this case means clear and precise rules that, by taking into account the technical characteristics of artificial intelligence-based solutions, create a regime that ensures complete respect for fundamental rights and rights created by legislative instruments, among them GDPR, without preventing or needlessly limiting the advancement of and deployment of AI systems. Thus, the author brings forth two hypotheses: 1) It is not possible to practice complete compliance with the requirement to ensure certain rights as they are guaranteed by the GDPR, especially by data processors employing complex machine learning systems that process vast amounts of data through complex multi-step sets of instructions; 2) It is possible to offset the possible unnecessary limitations on the deployment of AI systems by clarifying the definitions and reformulating certain provisions found in the GDPR.

In compiling this article the author aims to conduct interdisciplinary research connecting two domains – legal and technical. The primary method for the exploration of the hypotheses is the analysis of legal norms with reference to legal texts, academic literature from the legal field as well as academic literature from the technical field. Essentially, legal norms analysis shall constitute the primary approach to discussion of the subject matter, with academic literature supporting such analysis.

[18] Krakovsky, Marina. "Finally, a peek inside the 'black box' of machine learning systems". Retrieved from: https://engineering.stanford.edu/magazine/article/finally-peek-inside-black-box-machine-learning-systems. April 25 2019

[19] Pagallo, U. (2017). The legal challenges of big data: Putting secondary rules first in the field of EU data protection. European Data Protection Law Review (EDPL), 3(1). p. 39

1. KEY CONCEPTS

Before carrying out discussion of the subject matter of this paper, certain key concepts, both technical and legal, must be explained. The general theme of the research is artificial intelligence, or AI – a broad term that encompasses technological solutions ranging from product recommendations on entertainment websites[20] to algorithms facilitating high-frequency trading[21] to AI-powered virtual home assistants. According to the definition given by the UK Government Office for Science and mentioned in the introduction to this paper, AI "generally (…) refers to the analysis of data to model some aspect of the world." Thus, it is important to keep in mind when discussing the juridical aspects of AI that "intelligence" in this context does not carry the connotation that the term carries in day-to-day conversations – rather than describing a system that is capable of independent thought that mimics an intelligent human being, at the current state of AI advancement, the term should be seen as describing machines that are able to find patterns and draw conclusions, or make guesses, based on those patterns, hence "analysis of data to model some aspect of the world." While "AI" is used to refer to the more general field of intelligent systems engineering, the term "machine learning" can be defined as the current application of the various theories and practices developed by AI researchers[22]. Machine learning, defined in Intel's iQ as "the set of techniques and tools that allow computers to "think" by creating mathematical algorithms based on accumulated data"[23] is based on the practice of programming systems that are able to "learn" and change their behavior based on the data inputted into those systems; these systems may be unsupervised, meaning that the data given is not labelled and the system attempts to identify structures or objects based on the patterns in that unlabeled data, or supervised, meaning that certain inputs are labelled which enables the machine to create models of the world based on correctly labelled data, and then use and tweak those models when analyzing unlabeled data[24]. A further subset is "deep learning", which "involves feeding vast quantities of data through non-linear neural networks that classify the data

[20] Government Office for Science. Artificial intelligence: opportunities and implications for the future of decision making. 2015.
[21] *Ibid.*
[22] Marr, Bernard. What Is The Difference Between Artificial Intelligence And Machine Learning? Retrieved from: https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#b28d15f2742b. December 6 2019
[23] Landau, Deb Miller. Artificial Intelligence and Machine Learning: How Computers Learn. Retrieved from: https://iq.intel.com/artificial-intelligence-and-machine-learning/. August 17 2019
[24] UK ICO. Big data, artificial intelligence, machine learning and data protection. Retrieved from: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf. February 10 2019

based on the outputs from each successive layer."[25] Deep learning systems employ neurons (created to mimic the neurons present in living creatures) that are categorized in three groups, being input neurons, hidden layer neurons (where it is possible for multiple hidden layers to exist), and output neurons; after the input layer receives the data, these inputs are passed to the first hidden layer which then performs complicated mathematical computations, or arrives at some result through Boolean (true or false) functions, and passes on the results to the next hidden layer, with connections between the neurons in these layers being assigned weights, with the end result being contained in the output layer in the form of a prediction.

Regardless of which category of AI technology is applied, machine learning or deep learning, the training of these tools and their subsequent deployment by nature requires large amounts of data, the complexity of the processing of which makes it difficult for data processors to fulfil the requirement of transparency[26]. Another factor that requires attention is that in 2016 the amount of data generated by an average person every day was estimated at 600 to 700 megabytes, while by 2020 that number is expected to rise to 1.5 gigabytes – these amounts are an aggregate of all data-producing activities such as presence on social networks, browsing of the Internet, using of services that collect data for analytics or targeted advertising etc.[27] The processing of such vast amounts of personal information, which lacks in absolute transparency, and is facilitated through the use of automated systems thus limiting the scope of possible human intervention, is at risk of being at odds with some of the fundamental principles and provisions of the European Union General Data Protection Regulation. However, before initiating the discussion of the primary subject matter of the paper – that the European data protection regime in its current form affects data processors that employ automated system to process personal data and make decisions concerning data subjects in a way that is unclear – it is necessary to explore what "personal data" and "personal data processing" means within the context of the GDPR.

Regarding the former, according to the definition given in Article 4(1) GDPR the term "personal data" includes all information concerning an identifiable natural person, or the "data subject"; "identifiable natural person" is a person that can be identified by reference to certain characteristics that comprise that person's identity, such as name, location data, psychological

[25] *Ibid.*
[26] *Ibid.*, p. 19
[27] Landau, Deb Miller. Artificial Intelligence and Machine Learning: How Computers Learn. Retrieved from: https://iq.intel.com/artificial-intelligence-and-machine-learning/. August 17 2019

and genetic traits, economic qualities etc.[28] As for the latter, Article 4(2) GDPR defines "processing" as any operations that are performed on personal data, including the recording, storing, alteration, use, destruction etc. by automated or otherwise means.[29] Consequently, if a credit institution, when making decisions concerning loaning, employs an automated system that analyzes a person's income, spending habits, saving habits, and previous loan repayment history, and compares that data to similar data from other customers of that institution, such conduct indeed falls within the scope of the GDPR. In fact, section (4) of the same article devotes specific attention to "profiling", a term which applies more precisely to the many real-life applications of AI technology and defined as any form of automated processing conducted for the purpose of evaluating certain aspects or a particular state pertaining to a natural person, such as work performance, economic situation, interests, behavior etc.[30]; this definition is of particular interest for the purposes of this paper, as the focus will be on automated systems employed in the course of business. At the same time, some do not find this (or any) provision of the GDPR to govern AI specifically – Robert van Genderen argues that "Surprisingly, there is no vision on the use of AI and robotics in the GDPR. These concepts are nowhere to be found in the text or recitals."[31] This disagreement, and the general nature of the provision related to profiling, can be explained in technical terms. In order for an automated decision to be made, for example, regarding a person's creditworthiness, AI techniques under their strict definition need not be used; the difference between true artificial intelligence and "if statements" must be kept in mind. While creditworthiness assessments using complicated AI-systems provide for much greater accuracy, in a rudimentary manner they can be carried out in the following way (implemented in Python):

```
personal_income = float(input(""))
if personal_income < 1000:
        return False
else:
        return True
```

[28] Humerick, M. (2018). Taking AI personally: How the E.U. must learn to balance the interests of personal data privacy & artificial intelligence. Santa Clara High Technology Law Journal, 34(4). p. 402
[29] *Ibid.*
[30] *Ibid.*
[31] Ven den Hoven van Genderen, Robert. Privacy and Data Protection in the Age of Pervasive Technologies in AI and Robotics. European Data Protection Law Review vol. 3 issue 3. 2017. p. 346

This simple program takes the input from the computer operator in the form of a floating point number (numbers with fractional parts in case income is not a whole number, e.g. $857.50 per month) and assigns the input to the `personal_income` variable; then, it compares the income to 1000 (or any necessary value) and outputs whether the person is eligible for a loan in a `True/False` form. While any necessary number of comparisons can be programmed, it will be able to provide a very simplistic answer, and will require bank employees to constantly update the compared-to values, due to changes in the overall health of the economy, average salary levels, bank's resources, etc. Using AI, on the other hand, provides for a much more accurate result due to the ability of modern computers to process vast amounts of data of all kind in real time and to establish patterns between those pieces of data. The simple implementation would satisfy the characteristics of automated decision making under Art. 4(4) and would make the satisfaction of the requirement of explainability an easy task. However, an AI system is a much more complicated matter than a collection of if-statements and comparisons, and this complexity does not seem to be specifically addressed in the Regulation, bar the more general definition of "profiling".

As mentioned in the discussion above, the GDPR defines "data processing" as encompassing any operations performed on personal data; the definition includes a non-exhaustive list of operations that would qualify, among which are such actions as general "use", "adaptation", "alignment or combination"[32]. Machine learning systems employed in the course of analyzing data concerning natural persons use and combine that data in order to discover patterns necessary for making decisions or suggestions (e.g. to determine eligibility for loans[33]); these systems also use this data and the assessment of the correctness of prior conclusions and decisions for training and algorithm efficiency improvement purposes to provide more accurate results[34]. Thus, the subsequent operation of the machine is carried out taking advantage of the conclusions made through analyzing previously provided personal data, i.e. using of the data. This means that, while data points concerning a particular data subject may not be referred to by the system (for example, if processing was based on consent[35], which the data subject withdrew later on) the

[32] *Ibid.*
[33] UK ICO. Big data, artificial intelligence, machine learning and data protection. Retrieved from: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf. 2017. p. 40
[34] Google. Descending into ML: Training and Loss. Retrieved from: https://developers.google.com/machine-learning/crash-course/descending-into-ml/training-and-loss. December 4 2019
[35] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, O.J. L 119/36. §6(1)(a)

actual using of that data persists throughout the operation of the system. This point will be important in the discussion of the compatibility of employing AI-based systems and of the right to erasure, enshrined in Article 17 GDPR[36].

With these key concepts defined, it is possible to begin the exploration of the legal issues surrounding the deployment of automated data processing systems under the regime established by the EU GDPR.

[36] *Ibid.*, O.J. L 119/43

2. REQUIREMENT OF TRANSPARENCY

Article 12 of Chapter III of the GDPR, concerned with the rights of the data subject, establishes the principle of transparent information and communication; according to section 1 the data controller (defined in the regulation as the person or entity that determines the purposes and means of data processing, data processor is a separate term that describes the person carrying out processing on behalf of the controller[37]; one single person can qualify as both at the same time) must provide certain information, as stipulated in Articles 13, 14, 15, 22, and 34, to the data subject "in a concise, transparent, intelligible and easily accessible form, using clear and plain language"[38]. Under Article 13(2)(f), the controller, in order to ensure transparency, is to notify the data subject of the employment of automated decision-making and to provide information "about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject"[39]. According to Casey et al. "the protection appears to envisage a limited right for data subjects to understand and verify the basic functionality of certain automated decision-making systems."[40] The nuanced technical nature of machine learning technologies makes it difficult for data processors to provide transparent information concerning the logic behind these systems; the UK Information Commissioner's Office (hereinafter ICO) states in its report "Big data, artificial intelligence, machine learning and data protection" that "the complexity of big data analytics can mean that the processing is opaque to citizens and consumers whose data is being used"[41]; the entity has also mentioned that the requirement of outlining the logic behind processing in privacy notices has been argued against for the reason that "the analytics used in big data are too difficult to explain in terms that people can understand."[42] The ICO dismissed this argument by referring to the UK Data Protection Act (DPA) that requires privacy notices to explain the purposes of processing, and does not require notices to describe the technicalities of the employed algorithms[43]. However, in this dismissal the ICO seems to have disregarded the fact that while the DPA may stipulate that processors

[37] *Ibid.*, O.J. L 119/33
[38] *Ibid.*, O.J. L 119/39
[39] *Ibid.*, O.J. L 119/41
[40] Casey, B.; Farhangi, A.; Vogl, R. (2019). Rethinking explainable machines: The gdpr's right to explanation debate and the rise of algorithmic audits in enterprise. Berkeley Technology Law Journal, 34(1). p. 158
[41] UK ICO. Big data, artificial intelligence, machine learning and data protection. Retrieved from: https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf. February 10 2019. p. 27
[42] *Ibid.*, p. 62
[43] *Ibid.*, p. 66

should give explanations regarding purposes and not exact methods, EU regulations take precedence; the GDPR entitles a data subject to a transparent explanation of the logic involved, in clear and plain language. Nevertheless, the ICO agrees that big data analysis is complex and opaque; the report states that, in regards to certain services, the right to explanation/justification is likely to apply, and "In such circumstances, it may be difficult to provide a meaningful response to an individual exercising their right to an explanation (…) because, as Jenna Burrell notes in her paper on the opacity of machine learning algorithms, when computers learn and make decisions, they do so "without regard for human comprehension"[44]. Moreover, if an individual is willing to challenge a particular decision regarding him or her, the effectiveness of this challenge is debatable: "If a challenge were to be made, the complainant data subject would be challenging the accuracy of a present prediction made about their future behavior. (…) The complainant data subject would therefore need to challenge the basis on which the algorithmic logic was constructed (including the training data) rather than the application to the complainant in that particular case – but this challenge could not be based on the GDPR accuracy principle as it does not relate to the complainant data subject's personal data, but rather to the construction of the algorithm and the third party training data."[45] This way, not solely the decision, but the entirety of the system that brought it about must be scrutinized – a process which requires in-depth technical knowledge. Thus, the ICO rightfully considered this issue to be significant enough to have a section devoted to algorithmic transparency.

In discussing the maximizing of the transparency of the processes and reasoning behind suggestions or predictions output by a machine learning system, the ICO suggests algorithmic auditing as one approach to solving the problem.[46] Algorithmic auditability is the principle according to which developers of algorithms should make it possible for third parties to assess how a specific suggestion of output was reached – this approach enables transparency in explaining the logic and enables developers to trace outputs and find errors in training or the architecture of algorithms.[47] This goal is difficult to attain due to the complexity of these systems and their "black box" nature, factors which have not prevented the creation and proposal of different methods of tracing algorithms. Among these methods is LIME – Local Interpretable

[44] *Ibid.*, p. 54
[45] Butterworth, Michael. The ICO and artificial intelligence: The role of fairness in the GDPR framework. Computer Law & Security Review. Vol. 34, issue 2. April 2018.
[46] *Ibid.*, p. 86
[47] Diakopoulos, Nicholas; Friedler, Sorelle.How to Hold Algorithms Accountable. Retrieved from: https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/. November 17 2019

Model-agnostic Explanations – proposed by University of Washington researchers, which is "an algorithm that can explain the predictions of any classifier or regressor in a faithful way, by approximating it locally with an interpretable model."[48] This algorithm identifies the traits or data points that affected a particular prediction the most, so as to determine whether the machine learning model can be trusted – while a prediction may be correct in one particular case, the model may have reached that result in a faulty way. For example, when a model that is created to label images with dogs in them is given various images of dogs to analyze and then an unlabeled similar image to identify, the machine may label the image as one depicting a dog, however it might do so by noticing that the unlabeled input and some or majority of the training images have grass on them and determining that to be the key pattern; this means that instead of identifying the object of interest, the machine identified another irrelevant feature that was present on multiple occasions and guessed that feature to be the defining characteristic. LIME and similar approaches may indeed aid data processors in explaining to data subjects how exactly certain decisions were made by automated systems in regards to the traits or patterns found by those systems, however this may not shed any light on the actual logic involved, that is, why those exact traits or patterns matter. In addition to this, there are various technologies that implement machine learning principles with varying degrees of complexity – while approaches similar to LIME may suffice in providing transparency for less complex systems, it is possible they would not be able to explain logic and decision-making undertaken by truly big data systems. Van Genderen agrees, stating that "In particular, neural networks are often 'black boxes', in which the (decision-making) processes taking place can no longer be understood and for which there are no explanatory mechanisms."[49] Whether the inability to explain the operation of machine learning systems completely is an issue of fundamental importance in regards to the requirement for transparency is unclear. In a recent case[50] in the Netherlands concerning transparency in personal data processing, the lack of transparency – the inability to describe the process behind the reaching of a particular decision – has proven to be a factor. Ruling on the legality of the government's use of SyRI, an AI-based risk indication system employed by the Dutch government "to predict the likelihood of an individual committing benefit or tax fraud or

[48] Ribeiro, Marco Tulio; Singh, Sameer; Guestrin, Carlos. "Why Should I Trust You?" Explaining the Predictions of Any Classifier. Retrieved from: https://www.kdd.org/kdd2016/papers/files/rfp0573-ribeiroA.pdf. February 16 2019
[49] Ven den Hoven van Genderen, Robert. Privacy and Data Protection in the Age of Pervasive Technologies in AI and Robotics. European Data Protection Law Review vol. 3 issue 3. 2017. p. 348
[50] Henley, Jon; Booth, Robert. Welfare surveillance system violates human rights, Dutch court rules. Retrieved from: https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules. February 5 2020

violating labour laws"[51], and the adequacy of the accompanying legislation the Court expressed its opinion that "the SyRI legislation contained insufficient safeguards against privacy intrusions and criticised a "serious lack of transparency" about how it worked", particularly due to the algorithm the system uses being secret[52]. Additionally, by deploying the system in poor neighborhoods and in the absence of more information, the system may "amount to discrimination on the basis of socioeconomic or migrant status"[53], while regarding the "fair balance" required by the European Convention on Human Rights the court stated that the legislation failed to achieve such between the objective of preventing welfare fraud and the violation of privacy.[54] This court ruling shows the reluctance of courts operating within the European legal system to dismiss the principle of transparency in assessing the compatibility of employed technologies with privacy law. What is particularly important is that the SyRI system was used by a public body to combat fraud; in cases involving private entities employing AI-based data processing for financial benefit the judiciary is even less likely to be lenient.

[51] *Ibid.*
[52] *Ibid.*
[53] *Ibid.*
[54] *Ibid.*

3. COMPLIANCE WITH THE RIGHT TO ERASURE

An important right guaranteed by the GDPR is the right to erasure (or the right to be forgotten) enshrined in Article 17.[55] In its essence, this right provides a data subject with the ability to compel the data controller to erase personal data without undue delay. The compliance of data controllers and processors with this right is not complicated to achieve when the data is being used in a strictly structured and traceable manner, for example by an insurance provider that keeps client information concerning a natural person and uses it to process claims. In this case, if the person exits their contract with the provider and requests erasure, the provider simply needs to destroy all files relating to this person and not necessary for the fulfilment of contractual obligations with another person (e.g. if the two persons were parties to the same claim and only one exited their contract). The situation becomes less clear when the data concerned is used in a less straightforward and more opaque[56] way, by an AI-based system. This is explained by referring to the conclusion the author reached when discussing the definition of "data processing". As mentioned, the definition includes various operations that may be performed on the data, including general use, which becomes problematic when the learning nature of AI is taken into account. In order to be able to reach any result at first and then more accurate results, machine learning systems must be trained; the various training techniques are irrelevant in this discussion, as all techniques make use of data processing in slightly different ways. When a system is trained, the need to refer to the original data is not constant as it has already been used once for algorithm training; the original data is accessed when a direct referral to it is made, e.g. when a loan eligibility assessment tool is instructed to engage the particular subject's data in order to make a decision concerning him or her. When this referral is not made, the data stays idle in the system. A natural assumption would be that in order to comply with the right to erasure, the computer operator simply needs to delete the original data from the storage medium (when a bank client decides to sever their relations with a banking institution, the institution is no longer concerned with the possibility of that subject requesting a loan from that institution). This, however, does not account for the effect the initial insertion of that data has on the continued operation of the entire machine. In order the illustrate this, the various temporal points of operation can be regarded as states of the ML-based model, which are the collections of

[55] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, O.J. L 119/43

[56] Temme, M. (2017). Algorithms and transparency in view of the new general data protection regulation. European Data Protection Law Review (EDPL), 3(4). p. 477

usages of input data. During the initial training, the labelled input data is analyzed and manipulated by the model in order to find patterns enabling the model to output results relevant to the model's purpose – a film review classifier would search for words and phrases that hint at a review being positive or negative using the assigned labels and upon feeding of new unlabeled input data would use discovered patterns to label it. This new, machine-labeled data can then be used for re-training or continuous training, depending on the exact learning techniques used. Thus, upon each re-training the model would enter into a new state. As mentioned above, the initial data is no longer accessed by the model, however the patterns found due to the using of that data persist; this is problematic, as the right to erasure invokes an assumption that with the erasure the using of the data ceases and the data subject is "forgotten", which is not exactly the case in regards to the aforementioned states. Moreover, the object of training data, containing personal information, cannot be separated from the concept of the model that makes use of this training data – the accuracy of the model is directly influenced by the correctness and the extent of balance of training data. This, in turn, makes the economic benefit derived from the operation of the model directly dependent on personal data.

When an AI-based system is in operation, it continuously makes use of all data that has been inserted and that it was, and is being, trained on, including that data that has been removed from the device the model is run on but that has already gone through the model. In the context of personal data protection, this creates a situation where personal data does not cease to be exploited by the operator of the model for financial gain. Referring once again to the GDPR, in recital (7) it is stated that "Natural persons should have control of their own personal data".[57] This sentiment of control over ones' own data, coupled with the right to erasure and data portability, directs towards the identification of a quasi-property right over personal data without directly establishing an explicit property right. According to Boerding et al., although "the GDPR does not provide specific regulations for any form of data ownership"[58], the right of erasure which allows individuals to order data processors/controllers to destroy personal data without undue delay grants those individuals "a power of exclusive disposition concerning the processing of personal data that is (…) comparable with the power of the owner over his

[57] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, O.J. L 119/2

[58] Boerding, A.; Culik, N.; Doepke, C.; Hoeren, T.; Juelicher, T.; Roettgen, C.; Schoenfeld, M. V. (2018). Data ownership a property rights approach from european perspective. Journal of Civil Law Studies, 11(2), 323-370. p. 331

property"[59] while the principle of data portability "can be seen as another step towards a (privacy-based) concept of data ownership".[60] Boerding et al. also consider the requirement of free flow of personal data in the EU under the ePrivacy Directive to be a reference to the free movement of goods which "indicates that data is considered as a transferrable asset or at least as comparable to tangible goods".[61] Generally, information in the digital age is often seen as a commodity[62] that is traded by data holders to service providers in return for goods and services, e.g. where users may allow companies like Google to analyze the data contained in their emails in return for greater storage space.[63] The author of this article is of the opinion that with the entering into force of the GDPR, a strong case can be made regarding the establishment of data ownership enjoyed by natural persons of their personal data – indeed, the Regulation created a framework where individuals have control over how and for what purposes their data is processed and where data used by legal persons in the furtherance of their business-related interests is provided strong safeguards akin to those that exist to prevent bad faith exploitation of tangible assets.

Going back to the discussion of the effect of initial training data on the model as a whole, even when personal data is removed from storage mediums it, essentially, continues to be exploited by the model to make further predictions. This means that when an individual severs their ties with the data controller/processor and withdraws their consent in regards to the processing of their data, the model operator still continues to use that data for the furtherance of their interests. Thus, a situation is created where a regime of data ownership is created over personal data that carries with it certain property rights, which are not in this case satisfied (according to Boerding et al. the right to erasure "could be understood as a negative dimension of an exclusive right, i.e., the power to exclude others from using one's property"[64]). Even if the discussion departs from the argument that a regime of quasi-property rights is established over personal data, some scholars argue that inferences created "through deduction or reasoning rather than mere

[59] *Ibid.*
[60] *Ibid.*, p. 332
[61] *Ibid.*
[62] Prins, C. (2006). When personal data, behavior and virtual identities become commodity: Would property rights approach matter. SCRIPTed: Journal of Law, Technology and Society, 3(4), 270-303. p. 276
[63] *Ibid.*
[64] Boerding, A.; Culik, N.; Doepke, C.; Hoeren, T.; Juelicher, T.; Roettgen, C.; Schoenfeld, M. V. (2018). Data ownership a property rights approach from european perspective. Journal of Civil Law Studies, 11(2), 323-370.

observation or collection from the data subject"[65] should be seen as part of personal data granting those the safeguards contained in the GDPR.

3.1. Compliance with the right to erasure re computer storage

Additionally, the right enshrined in Article 17 must be examined in the context of the phrasing of its supplementary name – the right to be forgotten[66] – and of the way how computer memory functions. This concept has been brought to the forefront in the ECJ case C-131/12, where a Spanish citizen brought action against a Spanish newspaper and Google, Inc. arguing that the plaintiffs infringed on his right to privacy; the person had been a party to a bankruptcy auction where his home was repossessed, which was published as per Spanish law in the newspaper. When searching for the person's name on Google the results linked to this past incident. The ruling was in favor of the Spanish citizen, with the court deeming Google a data controller and affirming the right of data subjects to request search engine providers to remove results containing personal information.[67] Villaronga et al. argue that the application of this right to be forgotten clashes the intentions of the regulators with the complexity of technical environments[68]; in fact, they state that "While "data deletion" may seem to be a straightforward topic from the point of view of many regulators, this seemingly simple issue poses many practical problems in actual machine learning environments. In fact, "data deletion" requirements can be considered to actually border on the edge of impossibility."[69] This is due to the technical nuances of computer memory and databases functioning.

Databases are programs used for the structured storing and efficient provision of data, which is stored in files structured in mathematical tree form; trees are data structures that allow for fast searching and retrieval of information.[70] Modern databases must be consistent, durable, and allow for isolation of parallel transactions; in order to fulfil these requirements, databases possess certain features, such as the ability to audit transactions (controlling what data was

[65] Wachter, S.; Mittelstadt, B. (2019). right to reasonable inferences: Re-thinking data protection law in the age of big data and AI. Columbia Business Law Review, 2019(2). p. 515
[66] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, O.J. L 119/43
[67] Fosch Villaronga, Eduard and Kieseberg, Peter and Li, Tiffany, Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten (August 13, 2017). Computer Security & Law Review (Forthcoming). Available at SSRN: https://ssrn.com/abstract=3018186
[68] *Ibid.*, p. 6
[69] *Ibid.*, p. 2
[70] *Ibid.*, p. 9

changed at what time and by whom), intended rollbacks (in case of a crash, the database must store enough history on previous states to be able to roll back incomplete or corrupting transactions), and backups.[71] Thus, it is possible that any particular piece of data is stored at various database locations and backups. If, in order to fully comply with the right to be forgotten, information must be deleted permanently, the memory spaces containing that information must be identified and overwritten; however, in cases such as with database transaction logs, this "is especially impossible without seriously endangering the consistency of the database, or even simply breaking it altogether."[72]

When a deletion request is communicated to the database, instead of overwriting the data the space is marked as deleted and hidden from search indexes, as actual overwriting would seriously impact performance. The database searches for the relevant information proceeding through memory spaces until the one containing the information is found, the space is then marked for deletion, linked by the garbage offset (a collection of deleted and free space), and moved from the list of active records to the garbage offset-linked list, linking to itself and severing the connection to adjacent memory spaces. When additional spaces are marked to be deleted, they are linked into the deleted records list, and when a space is required to store a new record the garbage offset list is searched for suitable space to overwrite. A figure illustrating this process in the context of a single tree page is provided in the paper.[73]

From this it is evident that when the computer operator deletes data in a database, it is not actually removed but is instead hidden from being searched for, and may stay in its memory space for a long time until that space is reused. In AI context, the effects of data removal, in the general sense, has implications in regards to the operability of the system; research has been conducted into effects of deletion of single data points in systems that use pre-calculated results as reference data, extract patterns, and then use the learned rules on new data, with results fed back into the "knowledge base" in order to train for next runs[74] – while deletion of randomly-selected single data points did not have large-scale effects, this research does not account for real-life situations where people who want their data deleted might possess certain important commonalities[75], which might greatly affect the accuracy of subsequent results.

[71] *Ibid.*, p. 9-10
[72] *Ibid.*, p. 10
[73] *Ibid.*, p. 11
[74] Such algorithms require the different sets of data to be similar in terms of data structure and statistical properties.
[75] *Ibid.*

4. PROPOSALS

The issues of compatibility of the new European data protection regime with contemporary AI-based solutions and the direction the development of AI is undertaking seem to lie with ambiguity of certain key provisions of the GDPR when applied to these solutions. The requirement of transparency and the right of explanation in "clear and precise language" is easily satisfied where data processing is carried out in conventional manner – for example, storage of personal information in physical form by an insurer, which is accessed by a human when claims arise, is straightforwardly explainable to the data subject. However, when processing is as opaque and all-engulfing as it often is when AI-based solutions are introduced, it is near impossible to explain the procedure nor the extent of the processing in layman terms. This can be attributed to several factors.

First and foremost, in order to gain detailed understanding of any technology-related subject one should ideally possess a significant degree of knowledge of contemporary technologies, however, the description of the functioning of many daily-used technologies can be relayed without much field-specific jargon. For example, in order to make calls mobile devices make use of radio waves to transmit voice data over a network. AI on the other hand is often at the crossroads of multiple scientific fields – computer science, mathematics, statistics, neuroscience and, in some cases, even linguistics.[76] What complicates the matter furthermore is that even when the scientific domain of AI-based solutions is limited, the increasing complexity of those solutions in turn make explaining their functioning increasingly complicated, and even impossible (the "black box"[77] concept).

As it currently stands, the text of Article 15(1)(h) GDPR is as follows: "The data subject shall have the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data and the following information (…) (h) the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences

[76] Hu, S. (2019). Detecting Concealed Information in Text and Speech. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 402-412
[77] Krakovsky, Marina. "Finally, a peek inside the 'black box' of machine learning systems". Retrieved from: https://engineering.stanford.edu/magazine/article/finally-peek-inside-black-box-machine-learning-systems. April 25 2017

of such processing for the data subject."[78] As discussed in this paper, the possibility of provision of "meaningful information about the logic involved" is unlikely due to the complexity of that type of logic. In order to make the requirement of transparency more adequate in the context of AI-based solutions the Article could be amended as follows: "the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the attributes of the data subject that are most likely to influence or direct the process of automated decision-making, as well as the significance and the envisaged consequences of such processing for the data subject." Such an amendment would allow employees of data controllers, who do not possess specialized technical knowledge, to explain data processing in terms of most relevant characteristics rather than logic in general – e.g. in cases of automated loan offer approval technologies, the data subject would be informed that his or her income level, spending habits and saving habits are most likely to play a role in determining whether a loan shall be offered to him or her.

Regarding the right to erasure, Article 17 GDPR which establishes the right to erasure[79] does not specifically address the exercise of this right in the context of AI-based data processing. Due to the quasi-property right implications and the regime of, essentially, data ownership this creates lack of clarity, as discussed above. Training data cannot functionally, nor legally for this reason, be separated from the model as the accuracy and thus the economic benefit derived from the operation of the model is directly dependent on personal data that is used to train it. This juridical ambiguity can be avoided if a section is created that tackles the handling of trained and continuously operational models. A possible wording would be the following: "Where the controller makes use of automated decision-making based on artificial intelligence to process personal data, the request by the data subject for the erasure of personal data shall be considered satisfied when such personal data in readily accessible form is removed from the storage media it is stored on." Here, "readily accessible form" would refer to digital objects containing the personal data which can be viewed by persons without the use of specialized equipment or software upon being accessed – text files, spreadsheets etc. In order to address the peculiarities of computer memory and storage management (discussed in the section on databases and the

[78] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, O.J. L 119/46

[79] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), 2016, O.J. L 119/43-44

functioning of the mechanism of information deletion), an additional section may be added to Article 17 GDPR which may be worded as follows: “Upon the receipt of an erasure request, the erasure request shall be considered satisfied when the personal data is removed from the storage media it is stored on in a manner that does not allow reproduction without specialized tools.”

While amendments and additions of this sort may fill some gaps in legislation as it is applied to artificial intelligence-based solutions, this solution, considering the rate of advancement of AI and the ever-increasing interest in these technologies by economic actors, is likely to be temporary. In order to tackle these issues more efficiently the author is of the opinion that it is necessary to create a separate piece of legislation dedicated exclusively to regulating personal data processing by means of AI-based systems. This legislation should be created not solely by professionals in the legal field, but through cooperation with engineering and computer science experts in order to ensure that the provisions take into account the technical nature of the regulated matter to the fullest extent and provide the most efficient safeguards and regulations considering those technical peculiarities.

BIBLIOGRAPHY

Articles

1. Agata Ferretti, M. S. (2018). Machine Learning in Medicine: Opening the New Data Protection Black Box. *European Data Protection Law Review*, 320-332.

2. Amitai Etzioni, O. E. (2016). Keeping AI Legal. *Vanderbilt Journal of Entertainment & Technology Law*, 133-146.

3. Andrew D. Selbst, S. B. (2018). The Intuitive Appeal of Explainable Machines. *Fordham Law Review*, 1085-1140.

4. Boerding, A., Culik, N., Doepke, C., Hoeren, T., Juelicher, T., Roettgen, C., & Schoenfeld, M. V. (2018). Data Ownership - A Property Rights Approach from a European Perspective . *Journal of Civil Law Studies*, 323-370.

5. Butterworth, M. (2018, April). The ICO and artificial intelligence: The role of fairness in the GDPR framework. *Computer Law & Security*, pp. 257-268.

6. Casey, B., Farhangi, A., & Vogl, R. (2019). Rethinking explainable machines: The gdpr's right to explanation debate and the rise of algorithmic audits in enterprise. *Berkeley Technology Law Journal*, 143-188.

7. Diakopoulos, N., & Friedler, S. (2016, 11 17). *How to Hold Algorithms Accountable.* Retrieved from MIT Technology Review: https://www.technologyreview.com/s/602933/how-to-hold-algorithms-accountable/

8. Eleonore Pauwels, S. D. (2018). The Internet of Bodies: Life and Death in the Age of AI. *California Western Law Review*, 221-234.

9. Hu, S. (2019). Detecting Concealed Information in Text and Speech. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 402-412.

10. Humerick, M. (2018). Taking AI personally: How the E.U. must learn to balance the interests of personal data privacy & artificial intelligence. *Santa Clara High Technology Law Journal*, 393-418.

11. Katyal, S. K. (2019). Private Accountability in the Age of Artificial Intelligence. *UCLA Law Review*, 54-141.

12. Mazur, J. (2018). Right to Access Information as a Collective-Based Approach to the GDPR's Right to Explanation in European Law. *Erasmus Law Review*, 178-189.

13. Omer Tene, J. P. (2013). Judged by the Tin Man: Individual Rights in the Age of Big Data. *Journal on Telecommunications and High Technology Law*, 351-368.

14. Pagallo, U. (2017). The Legal Challenges of Big Data: Putting Secondary Rules First in the Field of EU Data Protection. *European Data Protection Law Review* , 36-46.

15. Prins, C. (2006). When Personal Data, Behavior and Virtual Identities Become a Commodity: Would a Property Rights Approach Matter . *SCRIPTed: A Journal of Law, Technology and Society*, 270-303.

16. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, 2 16). *"Why Should I Trust You?" Explaining the Predictions of Any Classifier.* Retrieved from https://www.kdd.org/kdd2016/papers/files/rfp0573-ribeiroA.pdf

17. Sandra Wachter, B. M. (2019). A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review*, 494-620.

18. Temme, M. (2017). Transparency in View of the New General Data Protection Regulation. *European Data Protection Law Review*, 473-485.

19. Tschider, C. A. (2018). Regulating the Internet of Things: Discrimination, Privacy, and Cybersecurity in the Artificial Intelligence Age. *Denver Law Review*, 87-144.

20. Tsesis, A. (2019). Data Subjects' Privacy Rights: Regulation of Personal Data Retention and Erasure. *University of Colorage Law Review*, 593-630.

21. Van den Hoven van Genderen, R. (2017). Privacy and Data Protection in the Age of Pervasive Technologies in AI and Robotics. *European Data Protection Law Review, 3*(3), 338-352.

22. Villaronga, E. F., Kieseberg, P., & Li, T. (2017, August). HUMANS FORGET, MACHINES REMEMBER: ARTIFICIAL INTELLIGENCE AND THE RIGHT TO BE FORGOTTEN. *Computer Law & Security Review*.

23. Zarsky, T. Z. (2017). Incompatible: The GDPR in the Age of Big Data. *Seton Hall Review*, 995-1020.

Legislation

24. European Commission. (2016, 5 4). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to

the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. *Official Journal*.

Web materials

25. Gokey, M. (2018, 11 20). *The best smart speaker you can buy: Amazon Echo vs. Google Home vs. Sonos One.* Retrieved from Business Insider: https://www.businessinsider.com/best-smart-speaker-amazon-echo

26. Joh Henley, R. B. (2020, 2 5). *The Guardian.* Retrieved from Welfare surveillance system violates human rights, Dutch court rules: https://www.theguardian.com/technology/2020/feb/05/welfare-surveillance-system-violates-human-rights-dutch-court-rules

27. Marr, B. (2016, 12 6). *What Is The Difference Between Artificial Intelligence And Machine Learning?* Retrieved from Forbes: https://www.forbes.com/sites/bernardmarr/2016/12/06/what-is-the-difference-between-artificial-intelligence-and-machine-learning/#b28d15f2742b

28. Marr, B. (2016, 12 8). *What Is The Difference Between Deep Learning, Machine Learning and AI?* Retrieved from Forbes: https://www.forbes.com/sites/bernardmarr/2016/12/08/what-is-the-difference-between-deep-learning-machine-learning-and-ai/#7eeda8f326cf

29. Tashea, J. (2017, 4 17). *COURTS ARE USING AI TO SENTENCE CRIMINALS. THAT MUST STOP NOW.* Retrieved from Wired: https://www.wired.com/2017/04/courts-using-ai-sentence-criminals-must-stop-now/

Web sources

30. Chris Smith, B. M. (2006, December). *The History of Artificial Intelligence.* Retrieved from University of Washington: https://courses.cs.washington.edu/courses/csep590/06au/projects/history-ai.pdf

31. EveryMac. (n.d.). *Apple Macintosh Plus (ED) Specs*. Retrieved from EveryMac: https://everymac.com/systems/apple/mac_classic/specs/mac_plus.html

32. Google. (2018, 12 4). *Descending into ML: Training and Loss.* Retrieved from Google Developers: https://developers.google.com/machine-learning/crash-course/descending-into-ml/training-and-loss

33. Krakovsky, M. (2017, 4 25). *Finally, a peek inside the 'black box' of machine learning systems.* Retrieved from Stanford Engineering: https://engineering.stanford.edu/magazine/article/finally-peek-inside-black-box-machine-learning-systems

34. Landau, D. M. (2016, 8 17). *Artificial Intelligence and Machine Learning: How Computers Learn.* Retrieved from Intel iQ: https://iq.intel.com/artificial-intelligence-and-machine-learning/

35. Macuga, T. (2017, 8 23). *What is Deep Learning and How Does It Work?* Retrieved from Australian Centre for Robotic Vision: https://resources.rvhub.org/what-is-deep-learning-and-how-does-it-work/

Reports

36. *Artificial intelligence: opportunities and implications for the future of decision making*. Government Office for Science. (2015).. Government Office for Science, London.

37. *Ethics Guidelines for Trustworthy AI.* Independent High-Level Expert Group on Artificial Intelligence. (2019). European Commission.

38. *Big data, artificial intelligence, machine learning and data protection.* Information Commissioner's Office. (2017).

Appendix 4. Non-exclusive licence

Non-exclusive licence for reproduction and for granting public access to the graduation thesis[1]

I _Aleksandr Kesa_____(author's name)

1.Give Tallinn University of Technology a permission (non-exclusive licence) to use free of charge my creation

Artificial Intelligence and the GDPR: Inevitable Nemeses?_____ ,
(title of the graduation thesis)

supervised by _____Tanel Kerikmäe_____ ,
(supervisor's name)

1.1. to reproduce with the purpose of keeping and publishing electronically, including for the purpose of supplementing the digital collection of TalTech library until the copyright expires;

1.2. to make available to the public through the web environment of Tallinn University of Technology, including through the digital collection of TalTech library until the copyright expires.

2. I am aware that the author also retains the rights provided in Section 1.

3. I confirm that by granting the non-exclusive licence no infringement is committed to the third persons' intellectual property rights or to the rights arising from the personal data protection act and other legislation.

_____

[1] *The non-exclusive licence is not valid during the access restriction period with the exception of the right of the university to reproduce the graduation thesis only for the purposes of preservation.*