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

Huu Dam Tran 195275HAJM

# GDPR-COMPLIANT AUTOMATED DECISION-MAKING: A WIDER DEFINITION FOR THE RIGHT TO OBTAIN HUMAN INTERVENTION AT AN EXAMPLE OF BASEL II ACCORD

Master's thesis

Supervisor: Thomas Hoffmann
Dr. iur., LL.M

Tallinn 2021

I declare that I have compiled the paper independently
and all works, important standpoints and data by other authors
have been properly referenced and the same paper
has not been previously been presented for grading.
The document length is 18692 words from the introduction to the end of summary.

Huu Dam Tran ……………………………
(signature, date)

Student code: 195275HAJM
Student e-mail address: thdam3012@gmail.com

Supervisor: Thomas Hoffmann
The paper conforms to requirements in force

……………………………………………
(signature, date)

Chairman of the Defence Committee: / to be added only for graduation theses /
Permitted to the defence
…………………………………
(name, signature, date)

## TABLE OF CONTENTS

# ABSTRACT

Nowadays, data has become the foundation, the raw material of modern technology, and the driver of growth and change. With the help of AI technology, data is utilized to automatedly make better decisions in almost every area, from private sectors to public ones. While this presents significant benefits, it can also bring concerns and problems from a legal perspective, particularly data protection issues. How to regulate these automated decision-making processes is of great importance. With this regard, the GDPR is a global standard on data protection rules that in principle, prohibits the use of these emerging technology on people. However, the exceptions of this become a rule with certain requirements and a set of additional safeguards. One of them is “the right to explanation”. There has been a huge debate in the academic of whether the GDPR implies the right to explanation. Some authors believe that the right to explanation does not exist because the regulation only requires an ex-ante explanation, while others admit there is one. Current research aims to analyze how to exercise effectively the existing safeguards against the use of automated decisions and propose improvements to ensure data subject’s privacy and legitimate rights. Qualitative analysis is conducted based on expert knowledge and relevant public documentation. The analysis showed that it should be called a right to meaningful information to end the battle and support a contestable framework that enables data subjects to understand and change the automated decisions, if applicable. Moreover, for further improvement, a wider definition of the right to obtain human intervention that takes human inputs in the different stages of system design, training, and testing should be adopted to help the GDPR achieves what it claims.

*Keywords: automated decision-making, the right to explanation, the right to obtain human intervention, a contestable framework*

# LIST OF ABBREVIATIONS

| | |
|---|---|
| AI | Artificial Intelligence |
| A-IRB | Advanced Internal Rating Based Approach |
| Art 29. WP | Article 29 Working Party |
| CJEU | the European Court of Justice |
| DPIA | Data Protection Impact Assessment |
| EAD | Exposure At Default |
| EDA | Exploratory Data Analysis |
| EU | European Union |
| F-IRB | Foundation Internal Rating Based Approach |
| GDPR | the European General Data Protection Regulation |
| LGD | Lost Given Default |
| PD | the probability of default |
| SA | Standardized Approach |

# 1 INTRODUCTION

“The world’s most valuable resource is no longer oil, but data”[1]. This is a statement in an article from The Economist newspaper that has been proven to be true in recent years. Ranging from small companies to big institutions, there has been a trend of utilizing data to enhance business efficiency. As the Internet network is designed to transmit data as quickly as possible and its infrastructure is interconnected around the world, data or digital information is constantly and massively generated online. Simultaneously, the Internet becomes a platform to obtain a huge amount of data. After the collection phase, data get analyzed and the process often involves AI algorithms without or limited humans’ inputs to extract business insights and make predictions. With a very high accuracy rate of predictions, many companies and institutions have been using this type of technology to make better decisions for business purposes such as hiring new employees, or inventory optimization and non-business purposes such as predicting cancers or fraud detections. And that is why it is called an automated decision-making process.

Not only the data economy demands a lot of initiatives to keep pace with the development of AI technology, but also data protection laws. Much of AI technology now is powered by machine learning algorithms. Concerns about back-box machine learning algorithms have an impact on data protection legislations, especially in the European Union with the adoption of the GDPR. In general, the automated processing of data to make significant decisions is prohibited by the GDPR but the exceptions have become new rules with associated safeguards. There has been a fierce debate of whether a ‘right to explanation’ for data subjects is created based on wordings or intentions of the regulation. Since the adoption of the GDPR, over the period of two years, countless industry leaders, media, researchers from different walks of life have contributed their unique perspectives to the subject. But the debate has largely focused on the ‘right to explanation’ and overlooked other fundamental rights, expressed explicitly in the GDPR that function as fundamental safeguards for protecting natural persons against harms from automated decision-making processes. Moreover, the effectiveness of the safeguards regulated in the GDPR on automated decision-making in practice also is in question.

[1] The Economist. n.d. The world’s most valuable resource is no longer oil, but data. Available at: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data> [Accessed 20 April 2020].

By analyzing legislations and academic commentary, this article proposes an alternative for the 'right of explanation' to end the battle of whether there is such a right and explain why this is just the beginning of a contestable framework to effectively exercise safeguards provided by the GDPR. Moreover, an example from the Basel II Acord of a wider definition of the right to obtain human intervention to implement humans' actions in different stages of automated processes or other words, to put human in the loop will be illustrated for further improvements to ensure data subjects' privacy and legitimate rights.

# 2 GDPR OVERVIEW

As technology developed and the Internet was invented, the EU recognized the need for data protection. In 1995, the EU passed the European Data Protection Directive, establishing minimum data privacy and security standards. In 2006, Facebook opened to the public. In 2011, a Google user sued the company for scanning her emails[2]. Europe's data protection authority realized the need for 'a comprehensive approach on personal data protection' and work to update the 1995 Directive[3].

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regards to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation or GDPR[4]), became applicable in May 2018. It is seen as the toughest privacy and security law in the world.

[2] ComputerWeekly.com. 2011. US woman sues Google over Gmail scanning. [online] Available at: <https://www.computerweekly.com/news/2240105327/US-woman-sues-Google-over-Gmail-scanning> [Accessed 11 May 2021].

[3] GDPR.eu. What is GDPR, the EU's new data protection law? - GDPR.eu. Available at: <https://gdpr.eu/what-is-gdpr> [Accessed 9 February 2021].

[4] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regards to the processing of personal data and on the free Government of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), [2016] OJ L 119/1.

The GDPR is now prime legislation of data protection in the EU. It stresses the EU's effort to make the law to cope with the changing society, technological developments and shows the world how the EU addresses the data protection issues.

## 2.1Data protection roles

The GDPR applies to four types of actors which are data subjects, data controllers, data processors, and data protection supervisory authorities (DPAs). The data subject is an identified or identifiable natural person, one who can be identified, directly or indirectly by name, an identification number, location data, or more factors specific to the physical, mental, cultural, or social identity of that natural person[5]. This person can be a legal person, and he or she must be the subject of at least one type of personal data processing by controllers or processers. A data controller is a person who decides why and how personal data will be processed. A Data processor is a person who processes personal data on behalf of a data controller. DPAs are independent public authorities that supervise, through investigative and corrective powers, the applications of the data protection law[6]. Since May 2018, they also govern how entities follow the GDPR norms and if they are in compliance with those new norms. DPAs also have the power to fine companies that breach data protection norms and GDPR norms. DPAs also provide expert advice on data protection issues and handle complaints.

For example, in the case of loan automated decision-making process, the data subject is the individual that applies for the loan, the data controller is the bank or the financial institution that offers the service of providing the personal loan, the data processor is the employee of the bank that directly handles the loan application and the supervisory authority is funded by the state to protect the fundamental rights and freedoms of natural persons concerning processing and to facilitate the free flow of personal data.

It should be noted that there is a case that two or more parties could be controllers of the same action of processing of personal data, so-called data joint controllership. While most guidance documents issued by EU regulators focus on the scenario where joint controllership arises from a legal arrangement between controllers, CJEU case law has broadened the definition by considering

[5] GDPR Art 4.1

[6] GDPR Art 51

situations where controllers are aligned merely by technical or organizational configurations[7]. Fashion ID is an online retailer whose website featured the Facebook ‘Like’ button that users can like the article and post on Facebook. Every time a visitor consults the website, his or her personal data such as name, IP address, cookies, etc. is transmitted to Facebook regardless of whether the visitor is a Facebook user or has clicked on the ‘Like’ button or not. In this case, the CJEU adopted a wide concept of joint controllership. It found that joint controllership can exist for specific states of the data processing, in this case, the initial collection of data and its transmission to Facebook. The court said that the website operator can qualify as a controller, joint with Facebook. Therefore, the CJEU concluded that as long as the website operator has a role in determining the purposes and means of the processing, it is a controller even if it doesn’t have access to the personal data collected and transmitted to the other party[8].

## 2.2Processing of personal data

The GDPR applies to personal data which means any information relating to an identified or identifiable natural person (‘data subject’). The GDPR stipulates that a natural person is identifiable when he or she “can be identified, directly or indirectly, by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person”[9].

‘Identification thus requires elements which describe a person in such a way that he or she is distinguishable from all other persons and recognizable as an individual.’[10]This definition provides for a wide range of personal identifiers to constitute personal data. Including name, identification number, location data, or online identifier, reflecting changes in technology and the way organizations collect information about people. Moreover, The GDPR applies to both automated personal data and to manual filling systems where personal data are accessible according to specific criteria. This could

[7] Jiahong Chen, Lilian Edwards, Lachlan Urquhart, Derek McAuley, 2020. Who is responsible for data processing in smart homes? Reconsidering joint controllership and the household exemption, International Data Privacy Law, Volume 10, Issue 4, November 2020, pp. 279–293.

[8] Case C- 40/17 - Fashion ID, ECLI:EU:C:2019:629.

[9] GDPR Art.4.1

[10] Kotschy, W., 2018. Handbook on European data protection law. Luxembourg: Publications Office of the European Union.

include chronologically ordered sets of manual records containing personal data. In the Nowak case, an exam candidate's written answer is the personal data of the candidate because it constitutes information that relates to the candidate and can be used to identify the candidate[11]. The court relies on these arguments. Firstly, the exam written answer's content reflects the extent of the candidate's knowledge and competence in certain fields. In the case of handwritten answers, it also contains information on the candidate's handwriting. Secondly, the purpose of this personal data is to evaluate the candidate's professional ability and his suitability to practice the profession concerned. Finally, the use of the information from the answer could have an impact on his or her rights and interests. For example, the grade of this answer's exam could increase or decrease the chance of entering the profession of the candidate. It is important to note that the assessment of whether personal data is accurate and complete must be made in the light of the purpose for which that data is collected. In this case, the answers are used to evaluate the professional ability and the competence of the candidate at the time of the examination. That ability is shown by any errors in those answers. As a result, such errors do not represent inaccuracy, which would give rise to the right of rectification.

On the other hand, the information that indirectly identifies a natural person is more difficult to be classified as personal data such as dynamic IP addresses. A dynamic IP address is the IP address that is dynamically allocated, each time the user connects to the network, his or her device is issued with a new IP address. Therefore, a dynamic IP address does not provide a website operator with sufficient information to directly identify the user, unless additional information is also available, such as information that the user provides when using the website. The CJEU stated that a dynamic IP address will be personal data in the hand of website operators if there is another party (such as an ISP) that can link the dynamic IP address to the identity the user or the website operator has a legal means of obtaining access to the information held by the ISP to identify the individual[12]. In general, if in the case of IP address could be used to identify the natural person, it can be identified as personal data.

Processing of personal data means any operation or set of operations that are performed on personal data or sets of personal data, whether by automated means. If you work with any data relating to a natural person then you process personal data, including collecting, recording, organization,

[11] Case C- 434/16 – Nowak, ECLI:EU:C:2017:994.
[12] Case C- 582/14 – Patrick Breyer v Germany, ECLI:EU:C:2016:779.

structuring, storage, retrieval, erasure or destruction, etc. For example, any organization using facial recognition technology that can recognize a face amongst a crowd then scan large databases of people to check for a match is the processing of personal data. Police can use this technology to identify individuals at risk or those who commit to criminal activities but the considerable increase of facial recognition technology has caused serious risks when it poses to the data protection the sensitive nature of the processing, the potential volume of people affected, and the level of intrusion to privacy it can create.

## 2.3Core principles

The GDPR has seven general data protection principles including fairness and lawfulness; purpose limitation; data minimization; accuracy; storage limitation; integrity and confidentiality; and accountability but data protection by design and default is at the core of the GDPR[13].

The principle of lawful, fair, and transparent requires personal information must be processed fairly, lawfully, and transparently[14]. This requires the controller to have at least one lawful basis for collecting and processing personal data. Article 6 of the GDPR provides six lawful grounds including consent, to carry out a contract, to meet a legal obligation of an organization, to protect vital interests of a person, for the performance of a task in the public interest, and the legitimate interest of a company or organization. No single basis is better than others, and there is no hierarchy among the six grounds[15]. Consent often referred to as 'opt-in' requires data subjects has consented to the processing of their personal data. For example, the online music providers must ask consent from the users to process their musical preferences to suggest. The second legal basis is that it is necessary to enter into performing a contract with the data subject, for example, online shops need to process data like name, delivery address, and card number of the customers to entering and performing a sale contract. The third legal reason is to meet a legal obligation of an organization, for example, the law obliges a company to provide personal data such as the weekly income of the employees to relevant authority

[13] Goddard, M., 2017. The EU General Data Protection Regulation (GDPR): European Regulation that has a Global Impact. International Journal of Market Research, 59(6), pp.703-705.

[14] GDPR Art 5.1(a)

[15] Gil González, E. and de Hert, P., 2019. Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles. ERA Forum, 19(4), pp.597-621.

to obtain social security cover. The fourth legal basis is to protect the vital interests of a person, for example where the hospital does not need a patient's consent to search his ID to check whether that person exists in the hospital's database to find previous medical history. The fifth legal ground is that it is necessary for the performance of tasks carried out by a public authority or private organization acting in the public interest, for example, a professional association such as a bar association vested with an official authority may carry out disciplinary procedures against some of their member. The last legal basis is for a legitimate interest of a company/ organization with the conditions of not contradict or harm the interests or rights and freedoms of other individuals, especially children.

The controller must also not do anything unlawful with personal data and must not process the personal data in a way that is unexpected, detrimental, or misleading to the data subject[16]. Moreover, the information and communication regarding personal data must be easily accessed and easily understood. For example, at the current state, many digital service providers display information about their processing of user's personal data in a so-called "Terms and Services" document. The principle of lawful, fair, and transparency requires those controllers to be clear, open, and honest about how and why they are processing personal data. The controllers must not hide any incidents related to their user's personal data.

The second principle is called "purpose limitation", according to which personal data can only be processed for the purpose for which it was obtained and not further processed in a manner that is incompatible with those purposes. The third principle is labeled 'data minimization' meaning the controller only processes the data required to meet the processing objective. In other words, processed personal data must be relevant and limited to achieve given processing purposes. Accuracy is the fourth principle requiring mechanisms need to be in place to detect and correct data errors.

The fifth principle, storage limitation requires data to be kept for no longer than required. This principle is encompassed in article 17 which formally introduces the so much debated "right to be forgotten"[17]. According to Article 17, individuals have the right to have personal data erased, aka the right to be forgotten. This right is not absolute and only applied in certain circumstances such as the

[16] GDPR Art 5.1.a.

[17] Politou, E., Michota, A., Alepis, E., Pocs, M. and Patsakis, C., 2018. Backups and the right to be forgotten in the GDPR: An uneasy relationship. Computer Law & Security Review, 34(6), pp.1247-1257.

personal data is no longer necessary with the purposes, the data subject withdraws his or her consent and there is no other legal basis or the personal data is processed unlawfully. The core objective of the right is to have unnecessary, unlawful data asked to be deleted by the data subject who is linked to that data. The CJEU in Google Spain had firmly established the significance of timeliness of personal data, stating that the search engine is a data controller and it must no longer make available to the public personal data if an easy search data is in a way that might be harmful to the data subject[18]. And at this time, the GDPR was not enforced yet as well as the right to be forgotten. Therefore, the court could not enforce a right that does not exist in the current legislation, instead, it used the application of the right to objection as a legal ground. The controller is not required to erase the personal data of the data subject but only not provide it to users upon web research using the name of the data subject as keywords. It seems that the court planted a seed, had sent a signal that it will recognize the essence of a right. And finally, the right to be forgotten was codified and to be found in the GDPR.

The sixth principle is integrity and confidentiality according to which personal Data must be kept safe, secure, and protected. This principle was introduced in the GDPR and did not exist in Directive 95/46/EC. The last and seventh principle is accountability requires the controller is responsible for and be able to demonstrate compliance by measures and/or records.

As mentioned, there are seven general data protection principles but data protection by design and default is at the core of the GDPR. According to Article 25 GDPR of data protection by design and by default, the controller needs to put in place technical and organizational measures to implement the data protection principle and safeguard individual rights. The concept is not new, previously known as privacy by design. It was a good practice but now, under GDPR, data protection by design and by default becomes a legal requirement. To implement data protection by design, the controller must put in place appropriate technical and organizational measures designed to implement data protection principles and integrate safeguards into your procession so that you meet the GDPR’s requirement and protect individual’s rights[19]. For example, the controller shall conduct a data protection impact assessment (DPIA) when the processing is likely to result in a high risk to the rights and freedoms of

[18] Case C- 131/12- Google Spain SL v. AEPD, ECLI:EU:C:2014:317.

[19] GDPR Art 25.

a natural person; implement technologies, processes to mitigate the risks that are discovered in DPIA; and have a privacy policy to be simple, easy to understand. To implement data protection by default, the controller must put in place technical and organizational measures to ensure only the process of personal data that is necessary for each specific purpose. To comply with this provision, the controller shall give individuals a simple, easy-to-access method for adjusting their privacy settings and exercising their data subjects' rights.

# 3 AUTOMATED DECISION-MAKING

## 3.1Automated decisions

In the current state, there is not any universal definition for an automated decision-making system. Narrowly, the European Commission described it as "decisions by technological means without human involvement"[20]. Therefore, automated decision-making could be simply understood as machines that can make a choice or a judgment about something by themselves. Historically, automated decision-making is originated coming from the evolution of decision-support systems that "are typically designed to help managers to report, analyze, and interpret data"[21]. This initial type of application's objective is to provide managers or executives with the information and the tools they need to make decisions. At the time, the role of human decision-makers was still crucial as they are actors responsible for making the final decisions while the decision-support systems simply support them with the decision as the name implies. While this initiative proved to increase the correctness and efficiency of a whole decision-making process it is believed that the results would be drastically improved if there is no human in the loop. Just take an example of a common computer function, copy and paste, how many times you see the function has performed wrongfully and how confident you are

[20] Art 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01).

[21] Davenport, T. and Harris, J., 2005. Automated decision making comes of age. Cambridge, MA: Sloan Management Review Association.

to believe you paste new text as exactly as the text you copied. As the result, automated decision-making systems, on the next level, were developed to minimize human involvement in the decision-making process as much as possible, some even without human intervention. Simply put, its purpose is to relieve humans from the task of making decisions. There are two types of automated systems ranging from traditional rules-based systems such as a system calculating a rate of payment following a formula set out in legislation to more complex systems such as those that use automated tools to predict and deliberate[22] with the use of big data and Artificial Intelligence.

### 3.1.1 Value of data

Data or digital data means any information that can be digitalized in other for a computer to work with.[23] That is a very broad definition for data, nowadays, as the advantage of technology, almost every information that can be digitalized ranging from the information of papers, pictures, sounds to even human emotions.

There are reasons why the Economist has a title of an article of "The world's most valuable resource is no longer oil, but data". It is worth comparing oil and data at the moment of this research. An oil refinery is an industrial cathedral, a place of power, drama, and dark recesses.[24] Without this type of resource, a lot of things existing in our modern life would disappear, whether from cars, trains, planes, or any other transportation to heating and electricity. Data is to this century what oil was to the last one: a driver of growth and change. Flows of data have created new infrastructure, new business, new monopolies, new politics, and crucially, new economics.[25]

What if there are too much data is generated? Data can be 'big' in two different ways. First, it can be 'big' due to the number of subjects, which may be individual people, animals, plants, objects, or social media activities. The other way that data can be 'big' by the number of characteristics or features on each subject.  For example, Facebook has been collected a huge amount of data from billions of its

[22] Commonwealth Ombudsman, 2020. Automated decision-making better practice guide. Available at: <https://www.ombudsman.gov.au/publications/better-practice-guides/automated-decision-guide> [Accessed 9 February 2021].
[23] BBC Bitesize. What is digital data? Available at: <https://www.bbc.co.uk/bitesize/topics/zj8xvcw/articles/zx3q7ty> [Accessed 9 February 2021].
[24] The Economist. Data Is Giving Rise To A New Economy. [online] Available at: <https://www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy> [Accessed 24 October 2020].
[25] Ibid.

users with all their features from the posts, likes, dislikes, comments of each subscriber. In 2020, 1.7MB of data is created every second by every person with a total of 2.5 quintillion bytes of data every day[26]. A quintillion is equal to 1 with 18 zeroes. Put simply, big data is larger, more complex data sets. These data sets are so voluminous that traditional data processing software just cannot manage them. To store and process such “big” data, there are various frameworks such as Hadoop, Storm, Hive, and Spark. Take Hadoop as an example, Hadoop uses a distributed file system known as Hadoop distributed file system[27] to store big data. How it works like the Internet breaking down big files into smaller chunks stored in various machines and make copies of it. That way, it can make sure that if one machine fails to do its task, the file is safe on another. The way it handles the processing task is the same, to process big data, the complex task is broken down into smaller tasks that will be executed by multiple machines simultaneously and then assemble the results at the end.

Many companies in the industry are already leveraging big data’s potential. It can be simply proven by looking at the largest companies in the world by revenue in 2019, it is a retail company Walmart with almost 524 billion dollars of profit. Walmart first started using historical sale data in 1990. Until now, data is wildly used to support business decision-making in a wide range of activities from store site selections, products, pricing to inventory and vendors. Amazon also lies in the top ten with over 280 billion dollars respectively. A very impressive figure for an e-commerce company 26 years old. Regarding profit, 8 out of 10 companies that are the most profitable in 2019 are generating revenue from the use of a massive database. Let us take a closer look into these titans, Apple, Alphabet (Google’s parent company), Microsoft, Amazon, and Facebook, they are the five most valuable firms in the world from Forbes’ annual list[28]. Data-driven startups also contribute to the new economy. They collect data, analyze it and turn it into novel clever services in a wide range of fields from camera tracking, eco-friendly infrastructure designs, to determine which drugs have a higher likelihood of success. More importantly, the value of data is increasing as those companies constantly apply new technology to exploit data. It would not be surprising if those companies will be even stronger in the

[26] TechJury. How Much Data Is Created Every Day In 2020? [You'll Be Shocked!]. [online] Available at: <https://techjury.net/blog/how-much-data-is-created-every-day/#gref> [Accessed 24 October 2020].

[27] Prathamesh Nimkar.2020. Hadoop Distributed File System.

[28] Forbes.com. 2020. The Worlds Most Valuable Brands. [online] Available at: <https://www.forbes.com/the-worlds-most-valuable-brands/#65aeede7119c> [Accessed 24 October 2020].

future. The sign of a data-driven economy is all over the place and has not been ever clearer. And another story behind these successes is the revolution of Artificial Intelligence.

### 3.1.2 Artificial Intelligence

The way data is utilized can be divided into 2 stages. First, the creation of a database derived from the collection of data from multiple sources, such as social media, purchase history, internet activities, etc. Second, the data is analyzed, and the process often involves with Artificial Intelligence algorithm to suggest recommendations or make predictions. As the application of AI in automated decision-making is widespread, as an extremely powerful technology, almost every tech companies or organizations are trying to utilize it, the author believes that it would be necessary to explore the nature of AI, understand the way it works before moving to the next chapters.

Artificial Intelligence is the study of intelligent behavior. Its goal is a theory of intelligence that accounts for the behavior of naturally occurring intelligent entities and that guides the creation of artificial entities capable of intelligent behavior.[29] In general, as The Council of Europe also states, AI could be distinguished as “strong” AI, with the ability to “contextualize very different specialized problems completely independently,” and “weak” to “moderate” AI, with the ability to “perform extremely well in their field of training”[30].

Weak AI is very common as it could be seen all around us. A few examples of Weak AI could be Google search, Siri or other personal assistants, self-driving cars. Such AI systems only focus on performing certain tasks, it is no doubt that they gain the most successful recognition of artificial intelligence today. Much of Weak AI now is powered by the massive improvement of machine learning and deep learning. The terms machine learning, deep learning, and AI are often used interchangeably but they are different to some extent, and distinguishing them is not an easy task. Venture capitalist Frank Chen provides a good insight to distinguish artificial intelligence, machine learning, and deep learning: “Artificial intelligence is a set of algorithms and intelligence to try to

[29] Martins, J., 1990. Michael R. Genesereth and Nils J. Nilsson. Logical foundations of artificial intelligence. Morgan Kaufmann Publishers, Los Altos, Calif., 1987, xviii + 405 pp. Journal of Symbolic Logic, 55(3), pp.1304-1307.

[30] Intelligence, A., AI, W. and Europe, C., 2020. Glossary. [online] Artificial Intelligence.

mimic human intelligence. Machine learning is one of them, and deep learning is one of those machine learning techniques."

Simply put, machine learning is a type of AI that feeds computer data and uses statistical techniques to find a model that fits the data as best as possible to predict outcomes without being programed explicitly. Machine learning consists of both supervised learning (using labeled data) and unsupervised learning (using unlabeled data). While deep learning is a type of machine learning technique to learn, that runs fed data through neutral network architecture, a computer system modeled on the human brain. This implies that any human contribution to the output of deep learning systems is "second degree"[31]. The neural networks contain various hidden layers through which the data is processed, allowing the machine to go "deep" in its learning, making connections, and weighting data input for the best results[32]. For example, Facebook uses machine learning algorithms to automatedly choose the contents suggested to users to keep them using the app as much as possible to increase the profit of the company via advertisements. These types of algorithms have no ethics or moral, they just try to maximize the time you spend on the social media site and such time directly bring the revenue to the company.

At the same time, Strong AI gains far less popularity. The creation of a machine with human-level intelligence that can perform any task is still controversial today and there is no strong evidence that such AI is generated successfully. As a result, data is utilized by companies with the support of AI to extract business values and automatedly make decisions. Then they use it to improve their business, attract more consumers who will generate more data, and the cycle restart. The more data gathers the more accurate decisions and the more business value.

[31] Gervais, D., 2019. Exploring the Interfaces Between Big Data and Intellectual Property Law. Journal of Intellectual Property, Information Technology and Electronic Commerce Law (JIPITEC) 22 Vanderbilt Law Research Paper No. 19-36.

[32] Builtin.com. 2020. What Is Artificial Intelligence? How Does AI Work? | Built In. [online] Available at: <https://builtin.com/artificial-intelligence> [Accessed 24 October 2020].

## 3.2 Application of automated processes

Automated decision-making is not a new concept. Australia has been at the forefront of this use of technology. Since the 1990s, they have been employing automated assistance – a rule-based system, to augment and automate administrative decision-making[33]. It changes the nature of administrative decision-making of the nation with over 20 provisions of Commonwealth legislation currently providing for decisions made by computer programs to be taken to be an official decision of a Minister, Secretary, or Regulator[34]. This is an example of traditional automated decision-making or traditional rules-based decision-making system which follows step by step following rules given by humans to achieve the goals.

On the other hand, unlike the traditional system, with the support of AI, modern automated decision-making systems do not necessarily follow explicit rules delegated by humans, rather based on the data and algorithms it has been trained on, the machine derives its own rules. Machine learning algorithms can be used in different ways in administrative decision-making. They can make decisions by themselves. A common example is the loan approval process. Nowadays, almost every bank or financial institution uses machine learning algorithms to predict the probability of default for loan applicants to decide whether the borrowers can return the loan or not. However, for the moment at least, it does not mean human involvement is entirely dismissed, humans can of course decide what decision-making to be automated and choose the data that the machine trained on.

[33] Swinson, J., Slate, R. and Fouracre, K., 2020. AI Guides | AI & Automated Decision Making. LEXOLOGY, [online] Available at: <https://www.lexology.com/library/detail.aspx?g=ffd70f05-d0f8-4dbc-b0c2-4395bf7265b9> [Accessed 10 October 2020].

[34] Ibid.

# 4 GDPR: COMPONENTS CONSTITUTE AUTOMATED DECISION - MAKING

In principle data subjects have “the right not to the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her"[35].

## 4.1 History

It is interesting when looking at how the Article developed through legislation procedure, evolving from focusing on a specific type of automated decision-making, Particularly, profiling to a broad notion. Beginning with Article 15 of the 1995 Data protection directive, every person not to be subject to a decision that “evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc.”[36]. At the time, there is no definition of profiling in the legislation. The provision only implied that the automated decision-making is limited to profiling that could be also defined as “the recording and analysis of a person's psychological and behavioral characteristics, to assess or predict their capabilities in a certain sphere or to assist in identifying a particular subgroup of people”[37].

The initial Commission’s proposal in 2012 for the GDPR took the same approach, the article was even title “Measure based on profiling” which is exclusively applied to the automated decision-making process of profiling. Particularly, Article 20(1) on measures based on profiling stipulates that “Every natural person shall have the right not to be subject to a measure which produces legal effects concerning this natural person or significantly affects this natural person, and *w*hich is based solely on automated processing intended to evaluate certain personal aspects relating to this natural person or to analyze or predict Particularly the natural person's performance at work, economic situation,

[35] Article 22(1) GDPR.

[36] Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.

[37] Khosrow-Pour, M., n.d. Advanced methodologies and technologies in modern education delivery. pp.585-594.

location, health, personal preferences, reliability or behavior.[38]" Moreover, this initial proposal contained a separate paragraph on the responsibility to inform the data subject about the existence of automated processing and 'the envisaged effects of such processing on the data subject"[39]. In the final version, the obligation to inform the data subject regarding such processing was moved to Articles 13 and 14 which require the controller to provide certain information to the data subject.

In the European Parliament resolution 2014, with the same patent, the article was again titled only to 'profiling' and there was a new right introduced that is the right to object profiling. Following that, for the first time, the definition of profiling is given: "means any form of automated processing of personal data intended to evaluate certain personal aspects relating to a natural person or to analyze or predict Particularly that natural person's performance at work, economic situation, location, health, personal preferences, reliability or behavior"[40]. This definition has not been changed too much until the GDPR. Interestingly, the European Parliament acknowledged the obligation of the controller to explain the decision reached. Article 20(5) on profiling states that "profiling which leads to measures producing legal effects concerning the data subject or does similarly significantly affect the interests, rights or freedoms of the concerned data subject shall not be based solely or predominantly on automated processing and shall include human assessment, including an explanation of the decision reached after such an assessment". This is one of the most important safeguards in the GDPR that will be explored in later parts of this paper.

Until position no 6/2016, the broad notion of automated decision-making is proposed by the council after a variety of discussions within the council and its preparatory and then is adopted by the European Commission and become article 22 in the GDPR. This is a significant improvement as it includes both types of automated decision-making. Since as discussed in Chapter I, there are two kinds of automated decision-making, which are the traditional rules-based decision-making system and the modern one. In the traditional decision-making system, it is not always necessary to relate to profiling.

[38] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regards to the processing of personal data and on the free movement of such data (General Data Protection Regulation)

[39] Article 20(4) proposal.

[40] European Parliament legislative resolution of 12 March 2014 on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regards to the processing of personal data and on the free movement of such data.

## 4.2 Main components

According to Article 22 GDPR, data subjects have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.

### 4.2.1 Nature of the processing

According to the Guideline on the GDPR of Article 29 data protection working party (Art. 29 WP), automated decision-making is the ability to make decisions by technological means without human involvement. It can be based on any type of data. Automated decision-making could be a machine algorithm automated deciding whether a customer's loan application is agreed by a bank or a financial institution and the result is automatically delivered to the individual, without any prior and meaningful assessment by a human. By forms, automated decision-makings could be divided into two categories, traditional automated decision-making, and AI-supported decision-making, which consists of profiling.

On the one hand, traditional automated decision-making, or traditional rules-based decision-making system is a system that follows step by step following rules given by humans to achieve the goals. A system automated calculates a rate of payment following a formula set out in legislation could be a common example.

On the other hand, profiling is presented as the most effective and wide-used technique in the application of machine learning algorithms. Article 4(4) of the GDPR defines profiling as any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, particularly to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location, or movements.

From the definition, three elements are consisting of profiling. Firstly, the processing must be in a form of automated processing, meaning there is no human intervention in the process.  From machine learning perspectives, this type of machine does not necessarily follow explicit rules delegated by humans, rather based on the data and algorithms it has been trained on, the machine derives its own rules. Secondly, the automated decision-making must be carried out on personal data or any information relating to an identified or identifiable natural people such as name, an identification number, location data, or more factors specific to the physical, mental, cultural, or social identity of that natural person. Thirdly, the objective of the profiling must be to evaluate the personal aspects of a natural person.

Profiling could be vividly explained by the algorithms Facebook has been created for its business purposes. According to The Social Dilemma, a docudrama film, the automated decision-making for Facebook users to be addicted to the app to increase the profit of the company[41]. The algorithm has no ethics or morals, they just try to increase the time users spend on the social media site as much as possible and such time spent on the platform will directly turn to the revenue of the company. These are millions of clusters of supercomputers running machine learning algorithms targeting each individual to figure out the best way to keep these individuals completely addicted and obsessed to use social media networks. These machine learning algorithms have access to thousands or millions of data points about users, all their interests and hobbies, and what is the emotion to press on them to get them to watch another video or read another article or post. This type of profiling is an extremely powerful weapon. As Article 29 stated, profiling means gathering information about an individuals or group of individuals and evaluating their characteristics or behavior patterns to place them into a certain category or group, particular to analyze and/or make predictions about, for example, their ability to perform a task; interests; or likely behavior[42].

Moreover, Article 22(1) GDPR refers to decisions ‘based solely’ on automated processing, meaning there is no human intervention in the decision process. This requirement is commonly understood as if the system takes any human review or opinion to reach its final decision, such a decision is not considered being based solely on automated processing. Or in other words, the decisive factor to

[41] The movie of The Social Dilemma.

[42] Art 29 Working Party, Guidelines on Automated individual decision-making and Profiling.

determine a decision is automated depends on whether human intervention can be embedded in the process and whether there is someone who has the authority and competence to change the decision. This does not mean that automated decisions are entirely devoid of human input[43]. At a moment, humans will decide what decisions are to be automated, the base system, or which datasets to train the algorithms.

### 4.2.2 Targets of the processing

Automated decision-making, including profiling, can be performed on a single individual or a group of individuals[44]. Individual profiling on a certain natural person to identify him or her and discovering characteristics, routines about him or her. On the other hand, group profiling deals with a set of people who share some common characteristics or patterns[45] - for instance, a political party, a group of people who live in a certain neighborhood, or the people who work in the same industry, etc. In which group profiles are far more common in the current economy, which is the result of three technical stages[46]. First, data is collected from a variety of sources including internal sources and external sources. The internal source is the database of the company containing the information about their customers such as their names, addresses, family sizes, which is the most often product they purchase or their financial situation including how often they use the services or buy the products, how much money they spend in one month. For example, Facebook has been collected a huge amount of data from billions of its users with all their features from the posts, likes, dislikes, comments of each subscriber.

On the other hand, data can also be obtained from the internet, an external source. At its most basic, the Internet is a global-scale network of computers to communicate altogether. As the internet network is designed to transmit the packets as quickly as possible and the Internet infrastructures are interconnected around the world and, it becomes a global marketplace where users can access infinite

[43] Swinson, J., Slate, R. and Fouracre, K., 2020. AI Guides | AI & Automated Decision Making. LEXOLOGY, [online] Available at: <https://www.lexology.com/library/detail.aspx?g=ffd70f05-d0f8-4dbc-b0c2-4395bf7265b9> [Accessed 10 October 2020].

[44] Pasquale, F., 2015. The Black Box Society: The Secret Algorithms that Control Money and Information. DigitalCommons@UM Carey Law.

[45] Gil González, E. and de Hert, P., 2019. Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles. ERA Forum, 19(4), pp.597-621.

[46] Privacy & Information Security Law Blog: UK ICO Issues Warning to Washington Post Over Cookie Consent Practices, November 21 (2018).

websites or applications. As a result, the Internet becomes the key-driven to generate and collect a massive amount of data. There is a wide-used technique to collect any public data from the Internet known as Web data scrapers. Secondly, the analysis of the data or exploratory data analysis (EDA) by statistical techniques to check the quality of data, figure out the patents of the data by checking its distribution, search for any correlations among different variables of specific profiles. The third stage involves applying several machine learning techniques to find a model that fits the data as best as possible and using that result for profiling other individuals.

### 4.2.3 Legal effects or similarly significantly affects

It is worth mentioning again that automated decision-making, including profiling, can have serious consequences for individuals. However, the GDPR does not define what is ‘legal’ effects concerning data subjects or ‘similarly significant’ affect him or her. According to Art. 29 WP, a legal effect requires that the decision affects someone’s “legal rights”, such as the freedom to associate with others, vote in an election, or take legal action[47]. This legal effect may also have an impact on a person’s legal status or their rights under a contract. While legal effects are easier to determine by examining the change or alteration of data subject’s rights, legal status, or legal duties, ‘significant’ effects are much vaguer[48]. Following the wording of the provision, generally, if a decision-making process does not affect people’s legal rights, the consequences of the decision could still fall within the scope of Article 22 if it produces an effect that is similar to that of a decision producing a legal effect. Some examples are given by Art. 29 WP such as decisions that affect a person’s financial status, their access to health services or education, etc.

## 4.3 Automated decision-making as authorized by the GDPR

As mentioned, data subjects have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her. However, the exceptions of this right themselves become a rule.

[47] WP29, Guidelines on Automated individual decision-making.
[48] Veale, M. and Edwards, L., 2017. Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling. Computer Law & Security Review 34(2) 2018, pp 398-404.

According to Article 22(2) GDPR, the prohibition "shall not apply if the decision (a) is necessary for entering into, or performance of a contract between the data subject and a data controller; (b) is authorized by Union or Member State law to which the controller is subject…; (c) is based on the data subject's explicit consent".

The first condition of automated decisions allowed by the GDPR is those that are "necessary" for entering into or perform a contract between the data subject and a data controller. This requirement can have different ways of interpretation. The WP 29 allows controllers to use automated means if it is necessary for "contractual purposes" and "pre-contractual processing"[49]. The requirement meets when there is no other "effective and less intrusive means to achieve the same goal exist" rather than automated decision-making processes[50]. The given example is fitting tens of thousands of applications for an open position[51]. This interpretation raises the question of whether an automated decision can ever be necessary when the same decision for entering and performing a contract can almost always be taken manually[52]. In the given example, humans may be capable to reduce tens of thousands of applications to a handful of them following some simple requirements. Then the question remains is that to which threshold a decision to enter and perform a contract is necessary to be handled by automated means.

Secondly, solely automated decision-making processes are authorized if they are authorized by Union or Member State law to which the controller is subject[53]. Such law must also regulate suitable measures to protect the data subject's rights and freedoms and legitimate interests[54]. Some examples for this are "fraud, tax-evasion monitoring and prevention purposes or to ensure the security and reliability of a service provided by the controller"[55]. However, currently, there is not yet any EU legislation that would specifically allow automated decision-making processes concerning Article 22(2)(b) GDPR.

[49] WP29, Guidelines on Automated individual decision-making, 23.
[50] Ibid.
[51] Ibid.
[52] Brkan, M., 2017. Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond, International Journal of Law and Information Technology, Volume 27, Issue 2, Summer 2019, pp. 91–121.
[53] GDPR 22(2)(b).
[54] Ibid.
[55] GDPR Recital 71.

Thirdly, the GDPR allows an automated decision if it is based on the explicit consent of the data subject. Explicit consent is not defined in the GDPR, rather according to it is explained by the WP29 on consent[56]. In which, explicit consent means that the data subject must give an express statement of consent, such as a signed written statement, filling out an e-form, or using an electronic signature[57]. With the popularity of online services where the data subject often is asked to give their consent online, it is not an easy task to clarify whether the obtained consent is explicated or not.

Another problem with online consent is that too often, the data subjects have no other options but to give consent to access or use certain online services. A potential Facebook user, for example, before successfully creating a Facebook account, is presented with an extensive list of clauses called terms of service and must, without any negotiation, either accept or decline the document to proceed further. A user gives his or her consent for the controller only by a simple click to the box of "I agree (with the terms of service)". It is worth noted again that it is not only easy but also there is no other way for a potential user but giving consent to using such a website or app.  This term of service can contain almost everything related to the business, including the services provided, how users' data will be used maybe include profiling, and the laws are to be applied when there is a dispute. The problem is that in general, people do not take time to read or do not have the knowledge to adequately understand such provisions. A Deloitte survey illustrated that 91% of people give their consent to such legal agreement without reading them, the figure jumps to 97% for those are in the ages of 18 and 34.[58] As a result, data subjects seemingly are forced to give their consent to profiling or even they are not even knowing about it to be able to use online services.

[56] Art 29 Working Party, Guidelines on consent under Regulation 2016/679, adopted on 28 November 2017; as last revised and adopted on 10 April 2018.
[57] Ibid, 18.
[58] Obar, J. and Oeldorf-Hirsch, A., 2016. The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services. Information, Communication & Society, pp. 1-20, 2018., TPRC 44: The 44th Research Conference on Communication, Information and Internet Policy, 2016.

# 5 GDPR: SAFEGUARDS OF AUTOMATED DECISION - MAKING

It is important to note that, whenever an automated decision is allowed, the data subject must be provided with appropriate safeguards. Based on the GDPR, the data subject has at least the right to obtain intervention, meaning that he or she can request that the decision is reviewed by a human, the right to express their point of view and contest the decision.

## 5.1 The battle of the right to explanation

### 5.1.1 Is there a right to explanation?

Articles 13(2)(f), 14(2)(g), and 15(1)(h), 22 of the GDPR mention the obligation of controllers to provide the data subjects the information about "the existence of automated decision-making, including profiling, …, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject"[59]. There has been a huge debate in the academic of whether the GDPR implies the right to explanation. This part of the paper will analyze both two points of view and explain the need to move on from this battle. It all started with a conference paper of Goodman and Flaxman[60]. The authors accepted the fact that there is a right to explanation in the GDPR without any explanation or any argument. They came ahead and prompted the question of what it means, and what is required, to explain an algorithm's decision.

#### *5.1.1.1 The right to explanation does not exist in the GDPR*[61]

On the one hand, Wachter et al. responded to this claim by saying a right to explanation of automated decision-making does not exist in the GDPR[62], including Articles 13(2)(f), 14(2)(g), and 15(1)(h), 22(3) of the GDPR. Particularly, they argued that the GDPR only requires an ex-ante explanation of how the system functions and not an ex-post explanation of the reasons behind the decision due to

[59] GDPR Articles 13(2)(f), 14(2)(g), and 15(1)(h)

[60] Goodman, B. and Flaxman, S., 2017. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". AI Magazine, 38(3), pp.50-57.

[61] This part will present main arguments of Wachter, S., Mittelstadt, B. and Floridi, L (2017).

[62] Wachter, S., Mittelstadt, B. and Floridi, L., 2017. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. International Data Privacy Law, 7(2), pp.76-99.

several reasons. Firstly, in Article 22 GDPR, a right to explanation is not mentioned. Instead, the article requires that after an automated decision has been made, the data subjects are granted additional safeguards to obtain human intervention, express his or her point of view or contest a decision but not to obtain an explanation for the reached decision. If yes, a right to explanation is only explicitly mentioned in Recital 71 which states that a data subject of automated decision-making "should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view**,** to obtain an explanation of the decision reached after such assessment and to challenge the decision"[63]. However, a Recital does not establish a legally binding right. As the European Court of Justice (CJEU) stated that: "Whilst a recital in the preamble to regulation may cast light on the interpretation to be given to a legal rule, it cannot in itself constitute such a rule.[64]" Or in other words, Recitals only guide how to interpret the Articles but are not themselves legally binding.

As a result, data subjects are not granted a legally binding ex-post right to explanation base on Article 22 GDPR. The safeguards indicated in Recital 71 are almost the same as those in Article 22(3) with the significant difference of a right 'to obtain an explanation of the decision reached after such assessment'. Secondly, speaking of Articles 13 and 14 GDPR, those provisions only specify notification duties of data controllers regarding the processing of personal data collected directly from the data subjects (Article 13) or a third party (Article 14). The claim of such Articles combined with Article 22(3) are legal bases for an ex-post right to explanation is mistaken because only the explanation of the automated decision-making system functionality is explicitly required by Article 13(2)(f) and 12(2)(g). Specifically, the data controller must inform the data subjects about: "the existence of automated decision-making, including profiling … [and provide data subjects with] meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing".

Moreover, the links of Articles 13(2)f and 14(2)g to the safeguards in Article 22(3) is not made in the GDPR. Rather, Articles 13(2)f and 14(2)g apply only to Articles 22(1) and 22(4), Article 22(3) only is featured by Recital 71. Therefore, the wordings in Articles 13(2)f and 14(2)g also support that only

[63] GDPR Recital 71.

[64] Case C- 215/88 - Casa Fleischhandels ECLI:EU:C:1989:331.

an *ex-ante* explanation is required before automated decision-making is given. Thirdly, Article 15(1)(h) is almost identical to Articles 13(2)(f) and Article 14(2)(h). The difference here is that the provision allows data subjects to actively access the information rather than waiting for the data controllers to inform them. Or Article 13 and 14 create notification duties for data controllers while Article 15 provides data subjects the right to access the information regarding the automated decision-making to examine the lawfulness of data processing and invoke legal remedies if needed. To conclude the position of there is no right of explanation of automated decision-making, Wachter et al. believe that the GDPR only requires an ex-ante explanation of how the system functions and not an ex-post explanation of how to arrive at the decision.

#### *5.1.1.2 The right to explanation does exist in the GDPR*[65]

On the other hand, Selbst and Powles support the existence of a right to an explanation based on Articles 13(2)(f), 14(2)(g), 15(1)(h), and 22(3) GDPR[66]. When a person is subjected to a decision based solely on automated processing that produces a legal effect or similarly significantly affects him or her, the GDPR grants the right to "meaningful information about the logic involved". Selbst and Powles make three reasons to explain why this could be seen as the right to explanation.

Firstly, since Articles 13 to 15 are all related to the rights of data subjects, the logic involved that data controllers provide should be meaningful to data subjects[67]. Or the provided information should be easy to understand by people without certain technical knowledge. Secondly, provided logic involved should be able to allow data subjects to exercise their legitimate legal rights such as the right to contest the decision as provided by Article 22(3) GDPR. And also, there should be a minimum threshold of functionality for the provided information, meaning the logic involved should be meaningful enough to facilitate that data subjects exercise their rights. Thirdly, the provision should be interpreted flexibly. Meaningful information should include an explanation of the principal factors that led to a decision that could be seen as a rigid way of defining a rule that may harm research and development. The authors gave an example of deep learning. It is true that the developers of some complex machine

[65] This part of the thesis will present main arguments of Selbst, A. and Powles, J.

[66] Selbst, A. and Powles, J., 2017. Meaningful information and the right to explanation. International Data Privacy Law, 7(4), pp.233-242.

[67] Kamarinou, Dimitra and Millard, Christopher and Singh, Jatinder, 2016. Machine Learning with Personal Data, Queen Mary School of Law Legal Studies Research Paper No. 247/2016.

learning systems such as neural nets even do not understand how the algorithms have made their decisions, it is impossible and redundant to force those people to come up with an explanation for reasons behind such decisions. Moreover, Article 13-15 GDPR states that, in addition to meaningful information, the data subjects must be able to know the 'significance and the envisaged consequences of such processing. The authors interpret this as information about how the results of the automated processing get used. For example, a bank would provide first meaningful information about the decision-making process itself and then the resulting downstream effects such as whether a loan will be granted at certain financial status or a certain interest rate. As a result, Selbst and Powles believe that the right to 'meaningful information about the logic involved' of automated decision-making is a right to explanation with a flexible interpretation that enables a data subject to exercise his or her right granted by the GDPR or other human right laws.

#### *5.1.1.3 A third alternative*[68]

Besides, a third alternative is supported by Maja Brkan. Not the right to explanation, the author admits that the GDPR grants the data subjects the right to be informed about the reasons for the automated decision[69] based on several observations.

Firstly, from the analysis of Articles 13(2)(f), 14(2)(g), and 15(1)(h) GDPR, the data subject should be informed about 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject' after the automated decision is already taking place. Particularly, such provisions should be interpreted as the data controller needs to inform the data subject about the reasons behind the taken decisions. While Wachter et al. state that the notification obligation about the system functions in Articles 13(2)(f) and 14(2)(g) GDPR occurs before making decisions, Maja argues that that claim only bases on the text of the introductory sentence of Article 13(2) which requires the data controllers to provide the information to the data subject at the time when personal data are obtained.

[68] This part will present main arguments of Barkan (2017).

[69] Brkan, (2017) *supra nota* 45.

At the same time, a similar phrase cannot be found in Article 14(2) which makes this argument mistaken in the cases when personal data is not collected from the data subject. Furthermore, both Articles 13(2)(f) and 14(2)(g) mention ‘the existence of automating decision-making’, the word ‘existence’ implies that the automated decision is already taking place. In terms of Article 15(1)(h), this Article does not specify the time when the data controller must communicate necessary information to the data subject. Maja takes the position of assuming the data subject can access data and related information after the automated decision was taken.

Secondly, the conveyed information to the data subjects on the reasons for the taken decision must enable them to express their point of view and to contest the taken decision as regulated in Article 22(3) GDPR. The data subjects need to be able to understand why and how an automated decision is made to effectively contest such a decision.

Thirdly, in contrast to Wachter et al. who said the right to explanation in Recital 71 GDPR is not legally binding, Maja acknowledges that by drafting provision in such a way, the legislators seemed to intentionally leave the final choice on the existence of this right to the CJEU. Rejecting the existence of a right of a data subject just because Recitals are not legally binding is too formalistic. Rather, Recitals would play a role as means to resolves ambiguity resulting from a joint of several GDPR provisions[70], in this case, the joint of Articles 13(2)(f), 14(2)(g), and 15(1)(h) and 22(3) GDPR.

Fourthly, the safeguards of human intervention, expression of point of view, and contesting are not the only possible rights for data subjects regarding automated decision-making. Article 22(3) states that those are “at least” safeguards for data subjects. Thus, such an additional safeguard mentioned in Recital could ensure a high level of data protection will be highly accepted.

Fifthly, accepting the existence of a right to be informed would support the principle of lawful, fair, and transparent meaning personal information must be processed fairly, lawfully, and transparently. Sixthly, having the opportunity to understand the reasons behind a decision is necessary for the prevention of discrimination in automated decision-making.

[70] Tadas Klimas and Jurate Vaiciukaite`, 2008, The Law of Recitals in European Community Legislation, 15 ILSA Journal of International & Comparative Law 26.

### 5.1.2 The need to move on from the battle

From those contradicted arguments, it is hard to say there is or there is not a right to explanation, especially based on the words of the GDPR. It seems to me that this battle is endless, then it is left for the DPA to decide there should be an obligation of the controller to explain how a certain decision is reached. Therefore, the author would propose a fourth alternative: a right to meaningful information. So that it is good enough to move on from the battle and create a framework to effectively protect data subjects' privacy, data protection and enable them to exercise the rights recognized in the GDPR and human rights laws.

First, the name of the right is not always the most important. It is not reasonable to just focus on the name of a particular right when analyze there is such a right, the proper way would be vice versa that is looking at the nature of a particular right and only after that, naming the right. From my observation, the right to explanation that a lot of researchers are fighting about should be named the right to meaningful information. All Articles 13(2), 14(2)(g), 15(1)(h) GDPR specify that the data controllers must provide the data subjects with 'meaningful information about the logic involved. By naming the right to meaningful information, it is easy to understand how it works, to give meaningful information regarding automated decision-making. Also, if sticking to the wording from Articles of the GDPR, the debate of whether Recitals are legally binding can be put to rest.

Second, the right to explanation meaning the data subject should know about the reason why the decision is made. It is useless for data subject because normally if there is such a right, a data subject would only use the right when they do not agree with the made decision, this right could be used by a controller to justify their decision by explaining for example how their system works, how the decision is made, maybe even in a plain language but the important thing here is that the data subject cannot do anything about it, they simply have to accept that is the case for them. Besides, taking to account how an automated processing works, especially deep learning works, even the engineering do not understand how the decision is made, how on earth they can explain fully to the data subjects. It is a feasible way for them to choose an easy, plain language, but maybe not correct to explain to the data subject and get away with the unfair decisions.

Third, the algorithm at the current state is not 100% correct, this is the reason why there is always a probability for any prediction to happen. For example, if a person is considered to have a loan, and acceptance would be made if the prediction is a 70% chance that he or she will then pay it back successfully.

Fourth, most researchers, again, so focus on a right that is most likely vague and forget about those rights that practically give data subjects the power to have their justices, which are the right to obtain human intervention, the right to express his or her views and contest the decision. Therefore, I can conclude that the battle of whether there is a right to explanation should be concluded by accepting the right to meaningful information. How do this right in combination with other rights can create a contestable framework for any automated decision-making process, especially profiling will be explained in the next parts of this thesis.

## 5.2 The right to obtain human intervention

According to Article 22(3) GDPR, the data subjects have the right to obtain human intervention of the part of the controller. There is no definition of human intervention in the GDPR or the Guidelines on automated decision-making and profiling of Art 29. WP. However, Art 29. WP requires ‘any review must be carried out by someone who has the appropriate authority and capability to change the decision’. It can be interpreted that the right to obtain human intervention means that data subjects have the right to have the automated decision on them being reviewed by a human with the ability to change the decision. He or she can have their fully automated decision becomes non-automated through human intervention[71]. There are a few arguments regarding the practical difficulties for the human revising the decision such as a human with limited capacities of data analysis would not be able to make a decision as accurate as machines. However, I believe this right could make a difference in enhancing the data protection for data subjects as Art.29 WP also states that “human intervention is a key element”.

[71] Brkan, (2017) supra nota 45.

Firstly, assume there is a ‘perfect’ AI machine, meaning it is perfectly correct, taking to account how the prediction of AI works, it will never be correct. There would be always noise or what is called outliers in probability. It means AI systems make a prediction based on the majority of the probability of an event likely to happen. For example, before signing the contract to lend a person a loan, an AI system will collect data from multiple sources and make a prediction of 99 percent (a very high probability) that he or she will be able to return the payment. Let imagine the company uses the system to make predictions for 1000 of its customers, there will be 10 people whose predictions are poor. Should those people deserve a second chance for human intervention? That is the reason why some scientists call those machines are perfectly imperfect. However, indeed, human intervention should not reduce the development of machine learning algorithms. This can be ensured if the broad definition of the right to human intervention is applied which will be elaborated in the later chapter.

Second, there would be no perfect AI system, based on algorithms, machines take a huge amount of data as inputs, use it to find a model that fits the data as best as possible. And make predictions based on that. There are three key areas on why AI fails: implicit bias, poor data, and expectation[72]. The first issue is implicit bias. As the name implies, machine learning algorithms become smarter by learning from data. If those data are biased, the algorithms will learn those biases and thus introduce biases in making decisions. It is also the second challenge, how to collect clean and qualified data. The last issue is expectation. In the current stage, many people do not trust machines, especially in sensitive aspects such as finance or health. For example, most people still prefer to see human doctors rather than machines even though the misdiagnosis rate of AI is much lower than a human doctor[73].

## 5.3 A contestable framework

Not only the right to obtain human intervention, but the author believes that the combination of rights or a contestable framework of rights should be applied to effectively protect the data subjects’ legitimate rights against the use of automated means. Particularly, the right to meaningful information, the right to express point of view, the right to contest the given decision, and the right to obtain human

[72] Sahota, N., 2020. Perfectly Imperfect: Coping With The ‘Flaws’ Of Artificial Intelligence (AI). [online] Forbes. Available at: <https://www.forbes.com/sites/cognitiveworld/2020/06/15/perfectly-imperfect-coping-with-the-flaws-of-artificial-intelligence-ai/#b85f8a9663ee> [Accessed 10 February 2021].
[73] Ibid.

intervention should work as a contestable framework for data subjects to ultimately have a chance to understand and change the given decisions.

It is natural that first and foremost, data subjects need to have enough information regarding the use of automated decision-making processes on their personal data. This includes at least the existence of such processes as well as the envisaged consequences that enable data subjects to understand two things, their personal data has been processed by automated means and what are the best as well as the worst scenarios that could happen to them after the decisions. This right could be exercised together with the right to access meaning data subjects have the right to access such meaningful information. For example, a person wants to apply for a loan in a bank. He or she should be informed or access to the following information. The loan process will be carried out by automated means, what types of data will be used to feed the algorithms and the outputs would be the applicant will get or will not get the loan without affecting their credit scores.

Then the data subject has the right to express his or her point of view. This right enables the data subject to have the means to communicate with the data controllers and the data controllers need to consider the opinions of the data subjects about the automated decision on them and have an obligation to respond to the data subjects. Following the previous example, when obtaining the result, the applicant should be able to speak for themselves, to provide or update their information, to explain their situation in detail, and hence, to ask the controller to change the decision after considering those data if they do not agree with the given decisions. The controller must reply to explain and answer whether they can change the decision without any serious review.

After that, if the data subject does not satisfy with the response, he or she could exercise the right to obtain human intervention, meaning the decision should be review by a person who has the appropriate authority and capability to change the decision. The reviewer should undertake a thorough assessment of all relevant data, including any additional information provided by the data subject to have a new decision. Interviews are usually necessary because human interaction remains important, as an actual person in some cases is far more informative than a mere collection of data related to him or her. As mentioned, the correct rate of algorithms can indeed be extremely high but it is still not absolute or 100%, there will always be some exceptions and this is the chance for humans to filter out those small number of cases.

There are a few best practices to facilitate the right to obtain human intervention for data subjects. While the new decision could be the same as the original one, it should be noted that all the rights mentioned above reside with data subjects, therefore, it should be no disadvantage for them after they exercise their rights. A solution for this would be, after exercising the right to obtain human intervention, it depends on the data subject to decide which decision is better for them, either the automated processing decision or human decision. In addition, the right to obtain human intervention should be the last means because it uses more resources and the right should be applied after the decision is made, otherwise, naturally, it is not an automated decision anymore and can harm overall technological development.

Finally, the data subject also has the right to contest the automated decision, to challenge the given decision. Therefore, it is expected that all those rights will be exercised together to create a system of contestable automated decision-making to enable data subjects to understand and change the automated decisions if applicable.

# 6 A WIDER DEFINITION OF THE RIGHT TO OBTAIN HUMAN INTERVENTION

## 6.1 The effectiveness of the current right to obtain human intervention

The right to obtain human intervention is a means to contest decisions that rely on an automated process. Accordingly, the subjects of automated decision-making have the right to have such decisions to be reviewed or changed by a person with proper authority. However, is this understanding or interpreting of the right effective in practice as it claims to be? There are three reasons why it is not the case.

First, currently, there is no limit for data subjects to exercise the right to obtain human intervention. Having enough inputs is one of the first and foremost conditions to create an algorithm that produces outputs as accurately as possible. Automated decision-making, such as profiling usually affects a broad range of natural people. Therefore, if a substantial number of individuals decide to exercise the right to obtain human intervention that involves humans to review the given decision, it may create a huge burden for the controller[74]. In terms of business efficiency, the burden of manual reviewing a significant amount of cases could slow down business processes, cost a fortune, and enormous human resources. Moreover, competitive companies could leverage the right to put in plenty of fake requests to damage the company.

The second question is raised as to whether a human with limited capacity could be able to review a machine decision properly. According to the WP29, the reviewer needs to have both 'authority and capability to change the decision' and need to make "a thorough assessment of all the relevant data"[75]. Despite the impressive complexity and processing power of the human brain, it is severely capacity limited[76]. Compared to computers that experience an increase in both processing power and memory power at an exponential level, it remains unclear how a human with a limited capacity of data analysis could be able to analyze and change a machine decision properly. Especially in the age of big data, usually, a huge chunk of data is used as inputs that make the task seems to be impossible for human brains to process.

Moreover, putting a human to the process could introduce biases, discriminations, or errors to the decision-making process. It is true that in this current state, automated decisions are also affected by discriminations when the historical data on which they are relied on, maybe be biased, incomplete, or

[74] Tad Hirsch, Kritzia Merced, Shrikanth Narayanan, Zac E Imel, and David C Atkins. 2017. Designing contestability: interaction design, machine learning, and mental health. In Proceedings of the 2017 Conference on Designing Interactive Systems. ACM, 95–99contestability: interaction design, machine learning, and mental health. In Proceedings of the 2017 Conference on Designing Interactive Systems. ACM, 95–99

[75] WP29, Guidelines on Automated individual decision-making, 27.

[76] Marois, R. and Ivanoff, J., 2005. Capacity limits of information processing in the brain. Trends in Cognitive Sciences, 9(6), pp.296-305.

even contain past discriminatory decisions[77]. For example, when a vision algorithm was trained to perform the task of distinguishing pictures of huskies from German Shepherds, it was very accurate at first, but it failed when predicting huskies that were kept as pets- it turns out that the algorithm discriminated the pictures based on identifying snow in the background[78]. However, in the future, it is possible to have machine learning algorithms that could be made to disregard discriminatory factors more effectively than humans[79].

Another great concern is that the current understanding of the right to obtain human intervention would slow down the development of automated decision-making technology. This issue is also shared by Hildebrandt who claims that data protection laws should ensure humans will not reduce the computable capacity of machines[80]. A great technology development does not happen overnight, rather it is a process of constantly improving over a long time. The same rule also applies to the improvement of automated decision-making. In a world in which every machine decision can be reviewed and changed manually would decrease the willingness to contribute of algorithm creators and machine learning engineers. Then it is useless to call "automated decision-making" when there is a human calling the final decision.

Moreover, trust also plays an important role to bring future technology into reality. To measure the success of a technology, calculate the popularity of the application of such an initiative is often calculated which has a strong positive correlation with the trust of users. For this reason, it is preferable to increase the accuracy and fair of the algorithms rather than allowing a human with the authority to review and change the outputs of such machine decision-making. In other words, the current legislation on the right to obtain human intervention on some level shows doubts over the correctness of automated decision-making and this could reduce the development of this technology.

[77] Žliobaitė, I., 2017. Measuring discrimination in algorithmic decision making. Data Mining and Knowledge Discovery, 31(4), pp.1060-1089.
[78] Spiegelhalter, D., 2020. The Art of Statistics Learning from Data. 1st ed. Penguin Random House UK.
[79] Kamarinou, Millard, and Jatinder Singh. 2016. Machine Learning with Personal Data. Queen Mary School of Law Legal Studies Research Paper 247. Queen Mary, University of London, United Kingdom.
[80] Mireille Hildebrandt. 2019. Privacy as protection of the incomputable self: from agnostic to agonistic machine learning. Theoretical Inquiries of Law, 20, 1.

## 6.2 Putting human in the loop

Because of the current ineffectiveness and difficulties of the current application of the right to obtain human intervention, a wider interpretation of this right was proposed by Lehr and Ohm.

Under a broader definition of the right to obtain human intervention, they claimed that human actions not only can be used to change the final decision that is the product of the automated processing but also be occur in the earlier stages of system design, training and testing[81] or in other words, putting human in the loop.

In the designing stage, machine learning engineers could decide what algorithms to be used and the categories of data used as inputs to train the algorithms. With the development of technology, machine learning algorithms could be applied to solve many problems, mainly regression and classification. The main task of the regression problem is to train a learner based on existing data and map the input to the corresponding output[82]. For example, from data of a customer as inputs, machines could predict this customer will buy or not buy outputs a particular product, outputs. There are several techniques to deal with this type of problem, namely Decision Trees, Linear Regression, or Random Forest. Classification problem as its name implies is a classification task[83] that could be solved using Kernel SVM, Decision Trees, or Logistic Regression. Among those techniques, machine learning developers are free to choose algorithms that can be explainable, such as Decision Tree and Logistic Regression to generate the outputs for an automated decision-making process. While other machine learning models are like black boxes, those algorithms are easy to interpret. For instance, the Decision Trees model provides a graphical and intuitive way to understand what the algorithm does[84]. Moreover, simplification is crucial when building a machine learning model as statistician George Box stated that "all models are wrong, but some are useful.[85]". Therefore, the algorithm designers need to decide

[81] David Lehr and Paul Ohm. 2017. Playing with the data: what legal scholars should learn about machine learning. UC Davis Law Review, 51, 653.

[82] Huang, J., Ko, K., Shu, M. and Hsu, B., 2019. Application and comparison of several machine learning algorithms and their integration models in regression problems. Neural Computing and Applications, 32(10), pp.5461-5469.

[83] Dokeroglu, T. and Sevinc, E., 2019. Evolutionary parallel extreme learning machines for the data classification problem. Computers & Industrial Engineering, 130, pp.237-249.

[84] Medium. n.d. Decision Trees Explained. [online] Available at: <https://towardsdatascience.com/decision-trees-explained-3ec41632ceb6> [Accessed 22 March 2021].

[85] So, R., 2017. "All Models Are Wrong". PMLA/Publications of the Modern Language Association of America, 132(3), pp.668-673.

how much uncertainty they are comfortable with and how much they are willing to accept to get a timely, explainable solution[86].

Sensitive data or data of a highly personal nature has been emphasized in the GDPR with additional protections such as different lawful bases for the processing[87], the obligation to carry out the data protection impact assessment[88]. According to article 9 GDPR, such data includes special categories of personal data such as information about an individual's political opinion, racial or ethnic origin, religious or philosophical beliefs, etc. These types of information should be considered by regulators of whether they could be used as inputs to train machine learning algorithms at the design phase. In addition, other types of sensitive data could also be personal data relating to criminal convictions or offenses as defined in Article 10 GDPR such as a hospital keeping patient's medical records or a private investigator keeping offenders' details or any information concerning the data subjects' performance at work, economic situation, health, personal preferences or interests, reliability or behavior, and location or movements[89].

In the training and testing stage, human intervention is not new. In practice, machine learning users often check and evaluate the predictions of the models and if the outputs are incorrect, they could tune the algorithms or re-verify the data and then again fed to the machine to make better decisions. The combination of human and machine intelligence that creates a continuous feedback loop allowing the algorithm to give every time better results is believed as the future of machine learning[90]. With this human interaction, it is not only easier for humans to understand how the algorithms work, but also the accuracy of the predictions is improved significantly.

However, it comes with a cost, often putting a human in a loop could decrease the overall predictive power and introduce more uncertainty. By empower human intelligence in an automated process, decision-makers or machine-learning developers would constantly face the tradeoff between the

[86] Vaughan, D., 2020. Analytical Skills for AI and Data Science. [S.l.]: O'Reilly Media, Inc.
[87] GDPR Article 9.
[88] Article 29 Working Party. 2017. Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is 'likely to Result in a High Risk' for the Purposes of Regulation 2016/679.
[89] Ibid.
[90] Medium. n.d. What is Human in the Loop Machine Learning: Why & How Used in AI?. [online] Available at: <https://medium.com/vsinghbisen/what-is-human-in-the-loop-machine-learning-why-how-used-in-ai-60c7b44eb2c0> [Accessed 22 March 2021].

transparency and the accuracy of a machine learning algorithm. Where automated decisions can have a major impact on people's lives such as deciding loan, insurance, the author believes the tradeoff is worth it.

Automated decision-making usually requires complex machine learning algorithms that seem impossible to explain in plain words. Lack of transparency is one of the reasons how the controllers can get away with the obligation of making sure data subjects understand how their information is processed fairly and accurately and prevent them to appeal against the given decisions. However, the demand for fairness, justice for critical automated decision-making required machine learning algorithms to be transparent, explainable, and able to be appealed against. By adopting a wider definition of the right to obtain human intervention meaning put human interaction in different stages of a machine learning process including design, training, and testing, it is feasible to build an automated decision-making system that is explainable to an average person. This type of explanation will not be necessarily explaining the algorithm itself but rather how decisions are reached. An example of an automated decision to decide whether a person will be granted a loan will be provided in the next chapter.

## 6.3 Human intervention at the example of Basel II Accord

This part of the paper will propose a sound data protection approach based on existing capital requirements for banks with the respect to the application of automated decision-making.

### 6.3.1 The importance of regulating automated processes in finance

No one in the current economy, especially in the financial sector, is immune from the myriad of the increasingly complex regulatory environment. The bank is no exception. Banks play a crucial role in our economy and society. In general, banking-related regulations are designed to fulfill two main roles. First, regulating banking operations, hence decrease risky financial behaviors and preserve the overall economic system. Second, baking-related regulations make sure the banking system is always in good health and maintain trust among clients.

Without proper legislation in place, many potential problems would surface. A notable example from the recent past would be the crisis enveloping global financial markets since August 2007 that was triggered by actual and prospective credit losses on US mortgages[91]. This is the most severe financial crisis from the time of the great depression in the 1930s. The crucial lesson to learn from this financial crisis is that the banks' and lenders' inability to calculate borrowers' probability of default or estimate credit risk can result in grave consequences for lenders and society in general. In which the probability of default means the likelihood of the borrowers will fail to pay back. Without a system to estimate the probability of default, the lenders were able to approve mortgages as many as they want to increase future revenue. As a result, high default rates and high mortgage approval rates are the main factors that led to this particular financial crisis in 2008.

The bankruptcy of big financial institutions led to a huge downturn in the economy. It affected every entity in the economy including the governments, companies, and individuals. In the European Union, the economic and financial crisis had generated a decrease in government revenues and an increase in government expenditures in terms of GDP[92]. And one of the reasons is that because Member States' governments had to bail those financial institutions out of the difficulty. Moreover, during the crisis, a huge number of firms had to be closed or had reduced their number of employees. The main way that firms have been negatively affected is because of the decreasing demand for goods and services[93]. Following that, thousands of jobs were lost along the way and millions of people experience financial difficulty. Lastly, overall labor productivity was also adversely affected by the crisis. Without sufficient motivation, it is generally the case that people do not work as hard as when the economy is growing at a normal pace. In the United Kingdom, over four years from 2008, the labor productivity is still below its previous peak level- Nicholas and María even proposed a hypothesis that economic capacity can be permanently damaged by financial crises[94].

[91] Luci Ellis, 2010. The Housing Meltdown: Why Did It Happen in the United States?, International Real Estate Review, Global Social Science Institute, vol. 13(3), pp. 351-394.

[92] Wahrig, L and Vallina, I., 2011. The effect of the economic and financial crisis on government revenue and expenditure. The European Union.

[93] Westergård-Nielsen, Niels C. & Neamtu, Ioana, 2012. How Are Firms Affected by the Crisis and How Do They React?, IZA Discussion Papers 6671, Institute of Labor Economics (IZA).

[94] Oulton, N. and Sebastia-Barriel, M., 2013. Long and Short-Term Effects of the Financial Crisis on Labour Productivity, Capital and Output. Bank of England Working Paper No. 470.

To avoid such great consequences, there are several regulations related to automated decision-making imposed on banks to maintain the stability of the overall economic system.

### 6.3.2 Basel II Accord

To prevent mentioned consequences, the bank regulators impose several requirements on banks, eliminating major risk for the stability of the economic system from the operation of banks. Particularly, the Basel Accords come up with three versions, in which Basel III is the third and the latest advancement, require the banks to obtain a certain amount of money as their capital so that they can cover the losses from the default of loans in their portfolios. The purpose of the system was to permit some flexibility in the allocation of capital, based on the perceived riskiness of various types of assets[95]. This requirement is also known as capital requirement. The Basel II Accord regulates how much capital the banks need to hold, how to define capital, and how to compare capital to risk-weighted assets. Particularly, the banks must hold at least 8% of the expected loss. This ratio is called the capital adequacy ratio meaning the ratio between the bank capital and risk-weighted assets. The assets are weighted according to three risks, including credit risk, market risk, and operational risk. The main advancement of the Basel II Accord compared to the Basel I is to ensure the banks' capital allocation is risk-sensitive, meaning the greater the risk a bank is exposed to, the greater the amount of money the bank needs to hold as capital to ensure the stability of the economic system[96].

After knowing the capital adequacy ratio, the next task is to calculate the necessary capital, the bank needs to find a way to estimate the expected loss or credit risk it is exposed to. It is a must for a financial institution to use an automated process that relies on machine learning techniques to estimate credit risk according to the regulations. The Basel II Accord has three pillars, namely minimum capital requirements, supervisory review, and market discipline. The first pillar – minimum capital requirements, presents the calculation of the total minimum capital requirements for three types of different risks, credit risk, market risk, and operational risk[97]. According to the Basel II Accord, the expected loss is a product of three components: the probability of default of borrowers (PD), the % of

[95] Wallison, P., 2009. CAUSE AND EFFECT: GOVERNMENT POLICIES AND THE FINANCIAL CRISIS. Critical Review, 21(2-3), pp.365-376.
[96] Shakdwipee, P. and Mehta, M., 2017. From Basel I to Basel II to Basel III. International Journal of New Technology and Research (IJNTR) ISSN:2454-4116, 3(1), pp.66-70.
[97] BCBS. 2005. Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Basel: Bank For International Settlements. Par 40.

exposure in case of default (LGD), and the exposure at default (EAD). The Basel II Accord prescribes that the regulators should allow banks to choose from three different approaches to calculating or modeling credit risk which is Standardized Approach (SA), Foundation Internal Rating Based (F-IRB) Approach and Advanced Internal Rating Based (A-IRB) Approach[98].

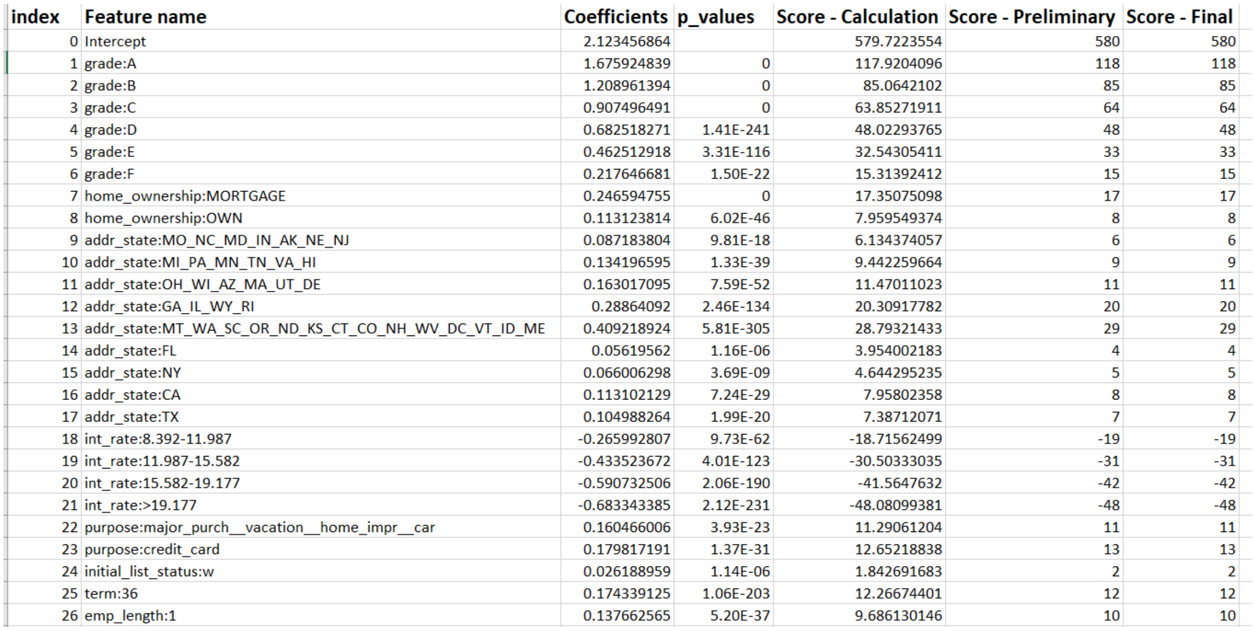
That is, they can choose one from three different approaches to calculate and model each of the three elements of the expected loss. For instance, in terms of where data is collected for modeling from, under the SA approach, the Basel II Accord requires the bank to use data from external credit agencies such as FICO credit scores for individuals and households to estimate PD, LGD, and EAD[99]. However, banks are also collecting a huge amount of data from their clients, hence acquire additional information about the behavior of applicants who are applying for loans. That information is extremely valuable to assess credit risk associated with each loan result in the better calculation of the required capital. This is the reason why under the other two approaches, regulators allow banks to estimate each of three components of the expected loss themselves. Particularly, under the F-IRB approach, banks can calculate PD from their data while LGD and EAD are given from external credit agencies and banks can model all the components by themselves under the A-IRB approach[100].

Moreover, it should be mentioned that the representation of credit risk information must also be easy to understand. The regulators require the model of the probability of default should allow an average person without any specific expertise to work with. There are a broad number of ways to present information about credit risk. When a credit risk agency carries out its evaluation, the credit risk of every entity including firms, countries, individuals, or households is presented by a credit rating. The higher an entity’s credit rating, the lower credit risk the bank is exposed to and that entity has the high creditworthiness. On the other hand, the lower an entity’s credit rating, the higher credit risk, and that entity have low creditworthiness. A person with a high credit rating will be more likely to be granted a loan compared to one with a lower credit rating.

In terms of individuals, credit rating is represented in the form of credit scores. The most popular credit score is the FICO score that is calculated and provided by a US company named FICO. The

[98] Ibid., Part 2.
[99] Ibid., Par 90.
[100] Ibid., Par 211.

FICO score ranges from 300 to 850, meaning the closer a person's credit score to 850, the higher his or her creditworthiness and the closer the credit score to 300, the lower their creditworthiness. To illustrate this idea in detail, the author conducted a complete credit risk project to create a scorecard and estimate expected loss for a bank based on Lending Club loan data from 2007 – 2018[101]. Linear regression and logistic regression are machine learning techniques to estimate each component of the expected lost equation[102]. This is an example of a scorecard used to calculate the FICO score for each loan application from the project.



| index | Feature name | Coefficients | p_values | Score - Calculation | Score - Preliminary | Score - Final |
|---|---|---|---|---|---|---|
| 0 | Intercept | 2.123456864 |  | 579.7223554 | 580 | 580 |
| 1 | grade:A | 1.675924839 | 0 | 117.9204096 | 118 | 118 |
| 2 | grade:B | 1.208961394 | 0 | 85.0642102 | 85 | 85 |
| 3 | grade:C | 0.907496491 | 0 | 63.85271911 | 64 | 64 |
| 4 | grade:D | 0.682518271 | 1.41E-241 | 48.02293765 | 48 | 48 |
| 5 | grade:E | 0.462512918 | 3.31E-116 | 32.54305411 | 33 | 33 |
| 6 | grade:F | 0.217646681 | 1.50E-22 | 15.31392412 | 15 | 15 |
| 7 | home_ownership:MORTGAGE | 0.246594755 | 0 | 17.35075098 | 17 | 17 |
| 8 | home_ownership:OWN | 0.113123814 | 6.02E-46 | 7.959549374 | 8 | 8 |
| 9 | addr_state:MO_NC_MD_IN_AK_NE_NJ | 0.087183804 | 9.81E-18 | 6.134374057 | 6 | 6 |
| 10 | addr_state:MI_PA_MN_TN_VA_HI | 0.134196595 | 1.33E-39 | 9.442259664 | 9 | 9 |
| 11 | addr_state:OH_WI_AZ_MA_UT_DE | 0.163017095 | 7.59E-52 | 11.47011023 | 11 | 11 |
| 12 | addr_state:GA_IL_WY_RI | 0.28864092 | 2.46E-134 | 20.30917782 | 20 | 20 |
| 13 | addr_state:MT_WA_SC_OR_ND_KS_CT_CO_NH_WV_DC_VT_ID_ME | 0.409218924 | 5.81E-305 | 28.79321433 | 29 | 29 |
| 14 | addr_state:FL | 0.05619562 | 1.16E-06 | 3.954002183 | 4 | 4 |
| 15 | addr_state:NY | 0.066006298 | 3.69E-09 | 4.644295235 | 5 | 5 |
| 16 | addr_state:CA | 0.113102129 | 7.24E-29 | 7.95802358 | 8 | 8 |
| 17 | addr_state:TX | 0.104988264 | 1.99E-20 | 7.38712071 | 7 | 7 |
| 18 | int_rate:8.392-11.987 | -0.265992807 | 9.73E-62 | -18.71562499 | -19 | -19 |
| 19 | int_rate:11.987-15.582 | -0.433523672 | 4.01E-123 | -30.50333035 | -31 | -31 |
| 20 | int_rate:15.582-19.177 | -0.590732506 | 2.06E-190 | -41.5647632 | -42 | -42 |
| 21 | int_rate:>19.177 | -0.683343385 | 2.12E-231 | -48.08099381 | -48 | -48 |
| 22 | purpose:major_purch__vacation__home_impr__car | 0.160466006 | 3.93E-23 | 11.29061204 | 11 | 11 |
| 23 | purpose:credit_card | 0.179817191 | 1.37E-31 | 12.65218838 | 13 | 13 |
| 24 | initial_list_status:w | 0.026188959 | 1.14E-06 | 1.842691683 | 2 | 2 |
| 25 | term:36 | 0.174339125 | 1.06E-203 | 12.26674401 | 12 | 12 |
| 26 | emp_length:1 | 0.137662565 | 5.20E-37 | 9.686130146 | 10 | 10 |

Two main columns that need attention are 'Feature name' and 'Score-Final' columns. The column of 'Feature name' depicts a variety of variables or features used to assess an applicant's credit risk such as credit grade, homeownership, the purpose of the loan, the length of employment, etc. The first feature, the intercept score is a special one. It is the score grated for every borrower at first. The column of 'Score-final' shows the respective credit score associated with each feature, respectively. This score can be positive such as the score of the feature 'grade: A' or it can be negative such as the 'int_rate:8.392-11.987' score. To sum up, initially, every borrower has the same score of 580. From

[101] https://github.com/DamHTran/Credit-risk-modeling

[102] Expected Lost (EL) = Probability of Default (PD) x Lost Given Default (LGD) x Exposure at Default (EAD).

there, depend on the different financial situation of the applicant, their FICO score will increase or decrease accordingly.

To give an example, a credit score for a borrower can be calculated as follow. Imagine an applicant who has a loan grade of B (85 scores), owns a house (8 scores), lives in New York (5 scores), his or her loan has an interest rate from 12 – 15.6% per year (-31 score) and he or she is applying for a loan with the term of 36 months (12 scores).  Take only these variables into account, the applicant will have the FICO score of 659 (= 580+ 85 + 8 + 5 – 31+12). This credit score is below the average score of US consumers, though many lenders will approve loans with this score. With this scorecard, data subjects without any financial expertise will be able to navigate themselves to calculate their credit score and have some idea about why their application is accepted or not.

### 6.3.3 Benefits of the wider definition of the right to obtain human intervention

The author believes that not only in finance but the same attention should also be paid to the general application of automated decision-making considering the effects those automated decisions can bring about.

For the banks, they would find it convenient, efficient to choose one from given approaches for modeling. It is true that with the advancement in the machine learning community, there are a lot of different sets of steps and available machine learning techniques to solve an existing problem that is proven to be much more accurate than traditional methods. At the same time, it could lead to confusion for decision-makers. With proper suggestions and requirements from regulations, the banks will be able to reduce a significant amount of their resources including money, human forces in finding and developing suitable machine learning models from scratch to handling the task at hand. The same pattern also applies to other critical areas such as healthcare, digital marketing, or transportation if applicable. For example, for cancer diagnosis, machine learning algorithms could be used to distinguish cancer and normal cell lines[103]. In the future when this type of automated process is

[103] Liao, Z., Li, D., Wang, X., Li, L. and Zou, Q., 2018. Cancer Diagnosis Through IsomiR Expression with Machine Learning Method. Current Bioinformatics, 13(1), pp.57-63.

prevalent, a hospital or a doctor would be extremely happy to adopt certain techniques that are given to them in the first place from the legislation.

Moreover, likely, the controller would also mention such regulations when obeying the obligation of providing meaningful information to data subjects. By proving compliance with regulations, the controllers of such automated decision-making could earn the trust of their customers and clients. Generally, the consumer’s trust is an important element of any company or institution’s objective because it directly generates revenue for the firm.

In terms of individuals, it is important to mention that the main objective to require human intervention in different stages including design, training, and testing of an automated decision-making process rather than the end phase of reviewing reached decisions by law is to protect legitimate rights of natural people against harmful impacts from such automated process. It is of course when narrowing down to a handful of machine learning techniques, regulators will choose the ones with the least harmful to data subjects as well as with the possibility of explanation. Prohibiting certain types of personal data used as inputs of algorithms also plays an important part in enhancing data protection for data subjects. Apart from that, data subjects will be more comfortable and confident when knowing the laws are already in place and the controllers successfully prove their compliance compared to dealing with some vague explanation based on black-box algorithms.

The adoption of the broader definition of the right to obtain human intervention helps to maintain the stability of the overall economic and social system when the application of automated decision-making based on machine learning algorithms becomes prevalent. As mentioned, the inability to control credit risk is one of the main reasons that led to global financial crises. The same pattern also applies to many other critical areas such as healthcare, transportation, or digital marketing. In a rapidly changing technological world, while creativity plays a crucial role in developing new initiatives, new methods to handle existing problems, when it comes to critical aspects in many people’s lives, there should be some principles, set of rules or framework to maintain the stability of the system.

As mentioned, this type of machine learning can be used as an extremely powerful weapon to manipulate people’s minds. These are supercomputers that contain millions of data points internet users including their interests, close friends, and what exactly impulse makes them behave as desired. For example, with access to millions of YouTube videos, Facebook posts, or Tweets posts, a machine learning algorithm can use that content to induce audiences. It can choose the most likely content that the target will click; it can prey on users' fear such as what age they will die, how bad the situation of Covid-19 is, or how fat they are right now; it can target consumers' emotions such as anger and anxiety by discussing about Black Lives Matter, mass shooting; or it can leverage users' friends, family, relationships to keep them continuously active on the Internet. Humans with the tendency to believe everything compatible with their views of the world or with their emotions until it is proven otherwise are very vulnerable to this type of weapon. The algorithms have no ethics or morals, they are created to carry out given tasks such as maximizing the time users spent on social media, thus increasing the company's revenue, broadcasting propaganda, interfering with an election. Indeed, the effects are not plain to be recognized immediately, but the consequences are truly serious.

However, the biggest challenge of this approach relies on the regulators. They need to obtain sufficient expertise in machine learning algorithms, data usage, and some certain professions. It is not an easy task. The Basel Accord for credit risk is a concrete example. It creates a framework for banks with different levels and ways of modeling credit risk depends on their capacity to ensure the accuracy of predictions, transparency of algorithms while providing enough flexibility for creativity and development. The same approach should be developed for other critical areas.

# 7. CONCLUSION

The sign of a data-driven economy is all over the place and has not been ever clearer. Data is utilized by companies and institutions with the support of AI to make better decisions. Much of AI technology now is powered by the massive improvement of machine learning and deep learning algorithms. Those automated means can be used to extract business insights, improve business optimization, and make decisions. The more data gathers, the more accurate the decisions, and the more business value.

At the same time, the popularity of technology brings certain concerns and problems into our daily basis. One of the most concerns is data protection. The EU has always insisted on the need to safeguard data protection and privacy. In the EU, the legal framework for data protection is concrete, especially the GDPR that plays a role as a landmark in the evolution of privacy. The regulation emphasizes the protection of individuals' rights by giving data subjects more control over their data. While the sole aim of the GDPR is to protect the data of EU citizens and residents, the impact of the GDPR goes far beyond the EU. Since coming into force two years ago from the time of this thesis, the EU's GDPR sets a global standard on data protection rules, considerably improves data protection in the EU and internationally.

According to the GDPR, in principle, data subjects have the right not to be subject to a decision which is based solely on automated processing, including profiling that produces legal effects or similarly significantly affects the data subjects. To be clarified as decisions based solely on automated means, the system must not take any human assessment to reach its final decision. However, this does not mean that automated processes are entirely devoid of human input. Humans, of course, can decide what kind of decisions are to be automated, the base systems, and which database to be trained. There are two types of automated decision-making processes which are traditional automated decision-making and AI-supported decision-making, including profiling. Profiling is the most common application of automated processes that uses personal data to evaluate certain personal aspects of a person to clarify them into a certain group to predict their ability to perform a task, behaviors, or interests.

Moreover, automated decision-making, including profiling must produce legal effects or similarly significantly affects the data subjects. As such automated processes can be performed on one single individual or group of individuals who share common behaviors or characteristics, it poses a great threat to individuals as well as society. While the "legal effect" requirement is easy to determine by examining the change of a data subject's rights, legal status, or legal duties, the "similarly significantly affects" condition is a lot harder to conclude and needs more guidelines. At the current state, decisions that affect a person's financial status, their access to health services or education are considered to produce similar effects to legal effects.

On the other hand, the exceptions of this right themselves become a rule. An automated decision is authorized by the GDPR if first, it is "necessary" for entering into or performance of a contract between the data subject and a data controller and there are no other alternative means that are effective and less intrusive to achieve the same goals. The remaining question is that to which threshold a manual solution is supposed to be not effective and more intrusive compared to automated means to prevent the abuse of data controllers. Second, if it is based on the explicit consent of the data subject. Explicit consent could be in a form of a signed written statement, filling out an e-form, or electronic signature. It is worth noting that too often online - services users seemingly are being forced to give their consent to profiling or even they are not even knowing about it in exchange to use such services. Not only it is convenient to click "I agree" but also there is no other way for them to communicate or negotiate the terms with the controllers.

Whenever an automated decision is authorized, certain safeguards must be provided to data subjects. There has been a huge debate in the academic of whether the GDPR implies the right to explanation. Some of them believe, based on the wordings of the GDPR, that the right to explanation does not exist because the regulation only requires an ex-ante explanation of how the system functions and not an ex-post explanation of how to arrive at the decision. While others support a flexible interpretation and admit there is a right to explanation that enables people without certain technical knowledge to understand the logic involved in the given automated decisions and allows them to exercise their rights granted from the GDPR and other human rights laws. The battle is endless, then it is left for the DPA to decide there should be an obligation of the controller to explain how a certain decision is reached.

Therefore, the author would propose an alternative: a right to meaningful information to resolve the concerns from both views, not only the name is compliant with the wordings of the GDPR, but also the content of the right enables data subjects to exercise their legitimate rights and support a contestable framework to effectively protect their privacy and benefits. Particularly, the right to meaningful information, the right to express point of view, the right to contest the given decision, and the right to obtain human intervention should work as a contestable framework for data subjects to ultimately have a chance to understand and change the given decisions while avoiding unnecessary burden on data controllers.

In which the right to obtain human intervention that allows data subjects to have the automated decision on them being reviewed by a human with the ability to change the decision could be seen as the key that determines the effectiveness of the framework. Machines are perfectly imperfect, there will be noises or outliers and biases to affect automated predictions. The objective of regulating the right is to give data controllers a second chance to review and improve given decisions by humans that have a significant impact on data subjects' lives. Moreover, the expectation of users or algorithm developers also proves an essential role to play of the right to obtain human intervention to protect data subjects' privacy and legitimate benefits against the use of automated processes.

Admitting the important intention behind regulating the right to obtain human intervention, in practice, there are still some concerns about whether it achieves what it claims. Those concerns include the scenario where a considerable number of data subjects simultaneously exercise their rights that may create a huge burden for automated means users, whether a human with limited capacity could be able to review a machine decision properly, the risk to introduce more biases, discriminations in the decision-making processes and the risk to reduce the development of technology. To resolve those concerns, the author supports the adoption of a wider definition of the right to obtain human intervention. In which, not only human intervention can be used to review and change the final decision, but it also can occur in the earlier stages of system design, training, and testing.

The Basel II Acord is an example of how to integrate human intervention in different stages of automated processes. After the financial crisis of 2008, regulators realized the demand for regulations of the automated decision-making processes, particularly for the capital requirement for financial institutions. Hence, the Basel Accords were adopted to ensure the stability of the overall economic system. According to the regulation, financial institutions can choose one from three approaches to design and implement automated decision-making processes including types of inputs, where to obtain data, the representation of credit risk information, and the used models to enable an average person to work with and estimate the minimum amount of necessary capital.

The same attention should be paid to other critical areas such as healthcare, transportation, digital marketing, or education if applicable. This is an improvement that is beneficial not only for individuals, institutions but also for social, economic systems. In terms of individuals, it increases trust among people who are subject to those types of automated processes. For the data controllers, they will find it easy and efficient to follow a given framework to design and utilize automated processes. Compliance with data protection laws will be also easier to prove. Lastly, the adoption of the broader definition of the right to obtain human intervention helps to maintain the stability of the overall economic and social system when the application of automated decision-making based on machine learning algorithms becomes prevalent. Then, the most challenge for regulators will be creating an automated decision-making framework that ensures the accuracy of the predictions, transparency of the algorithms while providing enough flexibility for creativity and development.

# REFERENCES

## Articles and books

Brkan, M., 2017. Do algorithms rule the world? Algorithmic decision-making and data protection in the framework of the GDPR and beyond, International Journal of Law and Information Technology, Volume 27, Issue 2, Summer 2019, pp. 91–121.

Jiahong Chen, Lilian Edwards, Lachlan Urquhart, Derek McAuley, 2020. Who is responsible for data processing in smart homes? Reconsidering joint controllership and the household exemption, International Data Privacy Law, Volume 10, Issue 4, November 2020, Pages 279–293.

Davenport, T. and Harris, J., 2005. Automated decision making comes of age. Cambridge, MA: Sloan Management Review Association.

David Lehr and Paul Ohm. 2017. Playing with the data: what legal scholars should learn about machine learning. UC Davis Law Review, 51, 653.

Dokeroglu, T. and Sevinc, E., 2019. Evolutionary parallel extreme learning machines for the data classification problem. Computers & Industrial Engineering, 130, pp.237-249.

Egger, M., Ley, M. and Hanke, S., 2019. Emotion Recognition from Physiological Signal Analysis: A Review. Electronic Notes in Theoretical Computer Science, 343, pp.35-55.

Gervais, D., 2019. Exploring the Interfaces Between Big Data and Intellectual Property Law. Journal of Intellectual Property, Information Technology and Electronic Commerce Law (JIPITEC) 22 Vanderbilt Law Research Paper No. 19-36.

Gil González, E. and de Hert, P., 2019. Understanding the legal provisions that allow processing and profiling of personal data—an analysis of GDPR provisions and principles. ERA Forum, 19(4), pp.597-621.

Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M. and Fdez-Riverola, F., 2013. Web scraping technologies in an API world. Briefings in Bioinformatics, 15(5), pp.788-797.

Goodman, B. and Flaxman, S., 2017. European Union Regulations on Algorithmic Decision-Making and a "Right to Explanation". AI Magazine, 38(3), pp.50-57.

Huang, J., Ko, K., Shu, M. and Hsu, B., 2019. Application and comparison of several machine learning algorithms and their integration models in regression problems. Neural Computing and Applications, 32(10), pp.5461-5469.

Kamarinou, Dimitra and Millard, Christopher and Singh, Jatinder, 2016. Machine Learning with Personal Data, Queen Mary School of Law Legal Studies Research Paper No. 247/2016.

Kamarinou, Millard, and Jatinder Singh. 2016. Machine Learning with Personal Data. Queen Mary School of Law Legal Studies Research Paper 247. Queen Mary, University of London, United Kingdom.

Khosrow-Pour, M., n.d. Advanced methodologies and technologies in modern education delivery. pp.585-594.

Liao, Z., Li, D., Wang, X., Li, L. and Zou, Q., 2018. Cancer Diagnosis Through IsomiR Expression with Machine Learning Method. Current Bioinformatics, 13(1), pp.57-63.

Luci Ellis, 2010. The Housing Meltdown: Why Did It Happen in the United States?, International Real Estate Review, Global Social Science Institute, vol. 13(3), pp. 351-394.

Marois, R. and Ivanoff, J., 2005. Capacity limits of information processing in the brain. Trends in Cognitive Sciences, 9(6), pp.296-305.

Martins, J., 1990. Michael R. Genesereth and Nils J. Nilsson. Logical foundations of artificial intelligence. Morgan Kaufmann Publishers, Los Altos, Calif., 1987, xviii + 405 pp. Journal of Symbolic Logic, 55(3), pp.1304-1307.

Mireille Hildebrandt. 2019. Privacy as protection of the incomputable self: from agnostic to agonistic machine learning. Theoretical Inquiries of Law, 20, 1.

Obar, J. and Oeldorf-Hirsch, A., 2016. The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services. Information, Communication & Society, pp. 1-20, 2018., TPRC 44: The 44th Research Conference on Communication, Information and Internet Policy, 2016.

Oulton, N. and Sebastia-Barriel, M., 2013. Long and Short-Term Effects of the Financial Crisis on Labour Productivity, Capital and Output. Bank of England Working Paper No. 470.

Pasquale, F., 2015. The Black Box Society: The Secret Algorithms that Control Money and Information. DigitalCommons@UM Carey Law.

Phan, N., Dou, D., Wang, H., Kil, D. and Piniewski, B., 2017. Ontology-based deep learning for human behavior prediction with explanations in health social networks. Information Sciences, 384, pp.298-313.

Politou, E., Michota, A., Alepis, E., Pocs, M. and Patsakis, C., 2018. Backups and the right to be forgotten in the GDPR: An uneasy relationship. Computer Law & Security Review, 34(6), pp.1247-1257.

Selbst, A. and Powles, J., 2017. Meaningful information and the right to explanation. International Data Privacy Law, 7(4), pp.233-242.

Shakdwipee, P. and Mehta, M., 2017. From Basel I to Basel II to Basel III. International Journal of New Technology and Research (IJNTR) ISSN:2454-4116, 3(1), pp.66-70.

So, R., 2017. “All Models Are Wrong”. PMLA/Publications of the Modern Language Association of America, 132(3), pp.668-673.

Spiegelhalter, D., 2020. The Art of Statistics Learning from Data. 1st ed. Penguin Random House UK.

Tad Hirsch, Kritzia Merced, Shrikanth Narayanan, Zac E Imel, and David C Atkins. 2017. Designing contestability: interaction design, machine learning, and mental health. In Proceedings of the 2017 Conference on Designing Interactive Systems. ACM, 95–99.

Tadas Klimas and Jurate Vaiciukaite`, 2008, The Law of Recitals in European Community Legislation, 15 ILSA Journal of International & Comparative Law 26.

Tsiamoulis, C., 2020. The impact of the principles of GDPR. Master. International Hellenic University. IHU.

Vaughan, D., 2020. Analytical Skills for AI and Data Science. [S.l.]: O'Reilly Media, Inc.

Veale, M. and Edwards, L., 2017. Clarity, Surprises, and Further Questions in the Article 29 Working Party Draft Guidance on Automated Decision-Making and Profiling. Computer Law & Security Review 34(2) 2018, pp 398-404.

Wachter, S., Mittelstadt, B. and Floridi, L., 2017. Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation. International Data Privacy Law, 7(2), pp.76-99.

Wahrig, L and Vallina, I., 2011. The effect of the economic and financial crisis on government revenue and expenditure. The European Union.

Wallison, P., 2009. CAUSE AND EFFECT: GOVERNMENT POLICIES AND THE FINANCIAL CRISIS. Critical Review, 21(2-3), pp.365-376.

Westergård-Nielsen, Niels C. & Neamtu, Ioana, 2012. How Are Firms Affected by the Crisis and How Do They React?, IZA Discussion Papers 6671, Institute of Labor Economics (IZA).

Žliobaitė, I., 2017. Measuring discrimination in algorithmic decision making. Data Mining and Knowledge Discovery, 31(4), pp.1060-1089.

Zuiderveen Borgesius, F., 2017. The Breyer Case of the Court of Justice of the European Union: IP Addresses and the Personal Data Definition. European Data Protection Law Review, 3(1), pp.130-137.

**Case law**

Case C- 131/12- Google Spain SL v. AEPD ECLI:EU:C:2014:317.

Case C- 215/88 - Casa Fleischhandels ECLI:EU:C:1989:331.

Case C- 582/14 - Patrick Breyer v Germany ECLI:EU:C:2016:779.

Case C-40/17 - Fashion ID ECLI:EU:C:2019:629.

Case C-434/16 - Nowak ECLI:EU:C:2017:994.

**Legislations and Guidelines**

Art 29 Working Party, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01).

Art 29 Working Party, Guidelines on consent under Regulation 2016/679, adopted on 28 November 2017; as last revised and adopted on 10 April 2018.

Art 29 Working Party. 2017. Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is ‘likely to Result in a High Risk’ for the Purposes of Regulation 2016/679.

BCBS. 2005. Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework. Basel: Bank For International Settlements.

Commonwealth Ombudsman, 2020. Automated decision-making better practice guide. Available at:<https://www.ombudsman.gov.au/publications/better-practice-guides/automated-decision-guide> [Accessed 9 February 2021].

Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data

European Parliament legislative resolution of 12 March 2014 on the proposal for a regulation of the European Parliament and of the Council on the protection of individuals with regards to the processing of personal data and on the free movement of such data.

Kotschy, W., 2018. Handbook on European data protection law. Luxembourg: Publications Office of the European Union.

Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the protection of individuals with regards to the processing of personal data and on the free movement of such data (General Data Protection Regulation)

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regards to the processing of personal data and on the free Government of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), [2016] OJ L 119/1.

**News**

Analytics Insight. 2021. Top 10 Countries Leading the Artificial Intelligence Race | Analytics Insight. [online] Available at: <https://www.analyticsinsight.net/top-10-countries-leading-the-artificial-intelligence-race/> [Accessed 10 February 2021].

Bbc.co.uk. What is digital data? Available at: <https://www.bbc.co.uk/bitesize/topics/zj8xvcw/articles/zx3q7ty> [Accessed 9 February 2021].

Builtin.com. 2020. What Is Artificial Intelligence? How Does AI Work? | Built In. Available at: <https://builtin.com/artificial-intelligence> [Accessed 24 October 2020].

Cassie Kozyrkov, 'Are you using the term 'AI' incorrectly?',Hackernoon (26 May 2018). [online] Available at: <https://hackernoon.com/areyou-using-the-term-ai-incorrectly-911ac23ab4f5> [Accessed 24 October 2020].

edpb.europa.eu. 2019. The Spanish Data Protection Authority fined the company Vueling for the cookie policy used on its website with 30,000 euros. [online] Available at: <https://edpb.europa.eu/news/national-news/2019/spanish-data-protection-authority-fined-company-vueling-cookie-policy-used_en> [Accessed 9 February 2021].

Forbes.com. 2020. The Worlds Most Valuable Brands. [online] Available at: <https://www.forbes.com/the-worlds-most-valuable-brands/#65aeede7119c> [Accessed 24 October 2020].

GDPR.eu. What is GDPR, the EU's new data protection law? - GDPR.eu. [online] Available at: <https://gdpr.eu/what-is-gdpr> [Accessed 9 February 2021].

internetlivestats. 2020. Twitter Usage Statistics. [online] Available at: <https://www.internetlivestats.com/twitter-statistics/> [Accessed 24 October 2020].

Medium. n.d. Decision Trees Explained. [online] Available at: <https://towardsdatascience.com/decision-trees-explained-3ec41632ceb6> [Accessed 22 March 2021].

Medium. n.d. What is Human in the Loop Machine Learning: Why & How Used in AI?. [online] Available at: <https://medium.com/vsinghbisen/what-is-human-in-the-loop-machine-learning-why-how-used-in-ai-60c7b44eb2c0> [Accessed 22 March 2021].

Prathamesh Nimkar. 2020. Hadoop Distributed File System. [online] Available at: <https://towardsdatascience.com/hadoop-distributed-file-system-b09946738555> [Accessed 24 October 2020].

Privacy & Information Security Law Blog: UK ICO Issues Warning to Washington Post Over Cookie Consent Practices, November 21 (2018).

Sahota, N., 2020. Perfectly Imperfect: Coping With The ‘Flaws’ Of Artificial Intelligence (AI). [online] Forbes. Available at: <https://www.forbes.com/sites/cognitiveworld/2020/06/15/perfectly-imperfect-coping-with-the-flaws-of-artificial-intelligence-ai/#b85f8a9663ee> [Accessed 10 February 2021].

Swinson, J., Slate, R. and Fouracre, K., 2020. AI Guides | AI & Automated Decision Making. LEXOLOGY, [online] Available at: <https://www.lexology.com/library/detail.aspx?g=ffd70f05-d0f8-4dbc-b0c2-4395bf7265b9> [Accessed 10 October 2020].

TechJury. How Much Data Is Created Every Day In 2020? [You'll Be Shocked!]. [online] Available at: <https://techjury.net/blog/how-much-data-is-created-every-day/#gref> [Accessed 24 October 2020].

The Economist. n.d. The world’s most valuable resource is no longer oil, but data. [online] Available at: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data> [Accessed 20 April 2020].

The Economist. Data Is Giving Rise To A New Economy. [online] Available at: <https://www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy> [Accessed 24 October 2020].

# APPENDICES

## Appendix 1. Non-Exclusive Licence

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

I Huu Dam Tran

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

GDPR-Compliant Automated Decision-Making: A Wider Definition For The Right To Obtain Human Intervention At An Example Of Basel II Accord

supervised by Thomas Hoffmann

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.*