Qualified Human Participation

How to Interpret "A Decision Based Solely on Automated Processing" for the Purpose of Art. 22(1) GDPR

MAYA RUESCH

TILBURG UNIVERSITY

Tilburg Law School
TILBURG UNIVERSITY
Tilburg, Netherlands 2019

Qualified Human Participation
How to Interpret "A Decision Based Solely on Automated Processing"
for the Purpose of Art. 22(1) GDPR

MAYA RUESCH

Student number:	U444253
Supervisors:	Merel Noorman
	Tom Chokrevski

Tilburg Law School
Tilburg University
5037 AB Tilburg
Netherlands

Table of Contents

List of Acronyms

ADM	–	Automated Decision Making
AWS	–	Autonomous Weapon System
CEV	–	Commission d'Examination des Voeux
CJEU	–	Court of Justice of the European Union
DPD	–	Data Protection Directive
EC	–	European Commission
EDPB	–	European Data Protection Board
GDPR	–	General Data Protection Regulation
ICO	–	Information Commissioner's Office
MHA	–	Meaningful Human Agency
MHC	–	Meaningful Human Control

Chapter 1

Introduction

1.1 Snapshot of Legal Outset

The most recent European Union (EU) legislation on the protection of personal data, the General Data Protection Regulation[1] (GDPR), contains several articles that regulate automated decision making (ADM). Among them is art. 22 GDPR, which constitutes the most elaborate article in relation to ADM.

In brief, art. 22(1) GDPR prohibits decisions based solely on automated processing in relation to a data subject, and art. 22(2) GDPR provides exceptions to this rule. It appears to be a clear set of provisions. When deliberating on art. 22(1) GDPR, however, the following questions soon arise: When is a decision based *solely* on automated processing? And what conditions does human conduct have to fulfil in order for a decision not to be based solely on automated processing?

The Article 29 Data Protection Working Party, which is the predecessor to the European Data Protection Board (EDPB), has published guidelines regarding art. 22 GDPR[2]. These guidelines were endorsed by the EDPB during its first plenary meeting, which renders them valid until future revisions by the EDPB are made.[3] The guidelines contain the following statements: "The controller cannot avoid the Article 22 provisions by fabricating human involvement."[4], and "To qualify as human involvement, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision."[5] As explained in Chapter 2, these statements provide little help to answer the questions stated above.

---

[1]Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1.

[2]Article 29 Working Party, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* (WP251, 2017) .

[3]European Data Protection Board, *Endorsement of GDPR WP29 Guidelines by the EDPB* (2018) ⟨https://edpb.europa.eu/news/news/2018/endorsement-gdpr-wp29-guidelines-edpb%5C_de⟩ accessed 2 May 2019.

[4]*WP251* (n 2) 21.

[5]Ibid 21.

The EDPB addresses the wording *a decision based solely on automated processing* with only a few sentences, and the threshold between a decision based solely or not solely on automated processing remains unclear. Since neither the GDPR, nor the EDPB, provide any clarifying information in this matter, this thesis attempts to bridge the interpretation gap by providing a set of requirements that explain when human participation suffices for the purpose of art. 22(1) GDPR. This thesis uses the term *qualified human participation* in order to refer to the degree of human participation that is required by art. 22(1) GDPR.

1.2   Consequences of Lack of Interpretation

Recital 10 to the GDPR states that, in order to establish a desirable level of protection for natural persons with regard to the processing of their personal data, harmonisation across Member States is necessary. However, coherent application of a legal article cannot be guaranteed if its provisions remain unclear. In order to resolve the ambiguity of the wording *a decision based solely on automated processing*, this thesis provides one possible interpretation of qualified human participation for the purpose of art. 22(1) GDPR.

The following example illustrates the result of a lack of consensus. In the case of credit-scoring, an ADM process is deployed in order to generate output that either suggests a grant or a denial of a person's credit application to a financial institution. Company A, B, and C engage in credit-scoring. In company A, the assessment of the creditworthiness of an individual is conducted by automated means only, i.e., no staff member checks the outcome. In company B, the scoring outcome of the ADM process, and an electronic profile that contains all credit-relevant information on the individual are sent to a staff member. This person reviews the outcome based on its coherence with the electronic profile. The staff member has the final say in whether the decision of the ADM process will be accepted, or revoked for a new decision. In company C, the outset is similar, but the relevant staff member, even though obliged to review and in the position to contest the output, has no doubt whatsoever regarding the adequacy of the output, and therefore always accepts the output.

The degree of human participation in these three scenarios varies significantly. While the practice of company B is likely to be permitted under art. 22(1) GDPR, it is unclear whether the practice of company A triggers the application of the article. The company A could argue that the fact that a human being has programmed the ADM process, and frequently reviews the algorithm, is sufficient to establish that a decision is not based solely on automated processing. Finally, in company C, human participation is intended, but not performed in an effective way.

In all these cases, the requirements of the EDPB that are stated in Section 1.1 have been adhered to. Human participation is not merely fabricated, and oversight can be considered meaningful, as the staff members are authorised to override the output, which means that all of the three scenarios could currently be compliant with art. 22(1) GDPR. As illustrated by this example, the lack of consensual interpretation regarding the wording *a decision based solely on automated processing* can lead to different understandings of the required degree of human participation. Companies

that apply ADM processes can adopt a level of human participation that primarily serves their own interests. Considerations such as cost effectiveness or operational simplicity can lead to a different degree of human participation than the legislator originally intended.

1.3   Thesis Approach

In order to solve the problems that are mentioned above, this thesis poses the following research questions: *How should the word solely in the context of art. 22(1) GDPR be interpreted? If solely is interpreted in a broad manner, how could the requirements for qualified human participation be defined?*

In order to answer the first research question, this thesis engages in doctrinal research on art. 22 GDPR. The wording *a decision based solely on automated processing* is analysed, contradictions of the provision identified, and clarifications to such contradictions are offered. Current official explanatory documents of the EDPB and the Information Commissioner's Office (ICO), and opinions in relevant literature are analysed. The results of this analysis lays the basis for the author's claim that the current state of interpretative guidance lacks substance and clarity. This motivates the direction of research of this thesis, which aims at providing a more detailed interpretation of qualified human participation.

In order to answer the second research questions, this thesis draws inspiration from human control concepts that are used in contexts outside the legal regime of data protection. These concepts are Meaningful Human Control[6] (MHC) and Meaningful Human Agency[7] (MHA).

MHC emerged from the debate on the deployment of fully autonomous weapon systems (AWS) in international conflicts. The concept conveys the idea that meaningful human control over the use of lethal force must be maintained, so that humans, and not computers, ultimately remain in control of, and thus morally responsible for, decisions about lethal military operations.[8] This thesis will focus on one account of MHC, which was proposed by Santoni de Sio and van de Hoven in their article "Meaningful Human Control over Autonomous Systems: A Philosophical Account".[9]

In "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems"[10], Ben Wagner elaborates on Meaningful Human Agency. In his paper, Wagner provides a list of requirements that aims to ensure meaningful agency when humans are part of an ADM process.

The objective of these two concepts, which is to ensure sufficient human control in an otherwise automated process, is in line with the rationale of art. 22 GDPR.

---

[6]Filippo Santoni de Sio and Jeroen van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account" (2018) 5 Frontiers in Robotics and AI 1.

[7]Ben Wagner, "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems" (2019) 11 Policy & Internet 104.

[8]"Killing by Machine: Key Issues for Understanding Meaningful Human Control" (*Article 36*, 2015) ⟨http://www.article36.org/wp-content/uploads/2013/06/KILLING_BY_MACHINE_6.4. 15.pdf⟩ accessed 20 May 2019.

[9]Santoni de Sio and van den Hoven (n 6).

[10]Wagner (n 7).

Therefore, the two sets of requirements of MHC and MHA are analysed, discussed, and compared as to their suitability to define qualified human participation. This thesis then combines the two concepts, and a new approach called Extended MHC is introduced, which combines the strengths of MHC and MHA. The Extended MHC criteria are then applied to a use case in order to illustrate their application to a GDPR setting. The use case in question is the French online platform Parcoursup, which is an online application that utilises personal data and ADM processes to match the wishes of a prospective student with available spots at higher education institutions.

1.4   Contributions

The main contributions of this thesis are:

- The identification of shortcomings of current explanatory resources regarding the term *a decision based solely on automated processing* for the purposes of art. 22(1) GDPR.

- An analysis how MHC and MHA could be used to define qualified human participation.

- A concept to define qualified human participation, called Extended MHC.

- An analysis of how Extended MHC applies to a specific case that is affected by art. 22(1) GDPR.

1.5   Limitations

This thesis will not provide a systemic analysis of art. 22(1) GDPR. This is due to two reasons. First, the objective of the thesis is to focus on the meaning of the specific formulation of *a decision based solely on automated processing* within the article, rather than providing yet another broad analysis of the article. The distinction between a decision based *solely* and *not solely* on automated processing is crucial in order to assess the application of art. 22 GDPR. If a decision process falls under the scope of art. 22(1) GDPR, such process is not permitted, with the exception that any of the exceptions in art. 22(2) GDPR apply. Second, an extensive examination would exceed the formal scope of this thesis. A limitation regarding Chapter 4 is that the goal of this thesis is not to identify strengths and weaknesses of the Parcoursup case in order to improve its process. The application of the Extended MHC criteria merely serves the purpose of demonstrating the application of the new concept to a GDPR-relevant setting.

1.6   Thesis Structure

Chapter 2 elaborates on the background and rationale of art. 22 GDPR, discusses key terms in relation to the provision, and provides an overview and analysis of the interpretation of *solely*. Thereby, the current lack of precise and workable explanation of the term is demonstrated. Chapter 3 lays out two human control concepts, MHC and MHA. The background of the concepts are explained briefly, and the requirements that flow from these concepts are explained, discussed, and finally compared. The most proficient conditions are identified, and combined to a new approach to define qualified human participation, called Extended MHC. Chapter 4 applies the Extended MHC criteria to the use case Parcoursup in order to demonstrate the applicability of the concept to a GDPR setting. Chapter 5 contains the major findings of the thesis, and will conclude the present work with suggestions for future work in relation to art. 22(1) GDPR.

Chapter 2

The Legal Outset:
Article 22(1) GDPR

The purpose of this chapter is three-pronged. First, it maps out the legal history of art. 22 GDPR, thereby providing the reader with a better understanding of the article. Second, it provides an overview of current official explanatory documents and views in legal literature on the meaning of *a decision based solely on automated processing*. Finally, it engages in an analysis of these sources. The results of this analysis show that the current state of interpretative guidance lacks substance and clarity. Thereby, the chapter motivates the present work's direction of research, which aims at providing a more detailed interpretation of qualified human participation.

Art. 22 GDPR constitutes the following:

Automated individual decision-making, including profiling

1. The data subject shall 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.

2. Paragraph 1 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 authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or

    (c) is based on the data subject's explicit consent.

3. In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.

4. Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.

2.1   Predecessor Article 15 DPD

The provisions set forth in art. 22 GDPR are not entirely novel, but were partly already laid down in art. 15 DPD. With effect from May 25, 2018, the GDPR repealed the 1995 Data Protection Directive[1] (DPD).[2] The 1978 French Act on Information Technology, Files and Personal Freedom and Various Provisions Concerning the Protection of Personal Data[3] is said to have inspired the EU legislator to lay down art. 15 DPD.

2.1.1   Rationale of Article 15 DPD

Art. 15 DPD constitutes the first legal provision addressing automated decision making in a data protection setting on EU level.[4] The article was angled at addressing, inter alia, concerns about automation in the early era of digitalisation. Such concerns were that decisions resulting from automation could be opaque and incontestable, and therefore unjustifiably preclude individuals from important services such as credit, housing, or insurance.[5]

In its 1990 DPD Proposal[6], the European Commission (EC) described its concern regarding the role that a human plays in an ADM process. The EC was alarmed by what it considered to be a decline of influence on the side of the human being in shaping important decisions in relation to himself. It deemed the use of individual data profiles as sole basis for a decision to be preventing the affected human to exercise any influence on the decision making process. Therefore, art. 15(1) DPD (in the 1990 DPD Proposal referred to as art. 14(1) DPD, constituting the version prior to art. 15(1) DPD) pursued to safeguard the participation of the data subject in the

---

[1]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 [1995] OJ L281/31.

[2]Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1, Art. 94(1).

[3]Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés .

[4]Isak Mendoza and Lee A Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling* (Springer 2017) 3.

[5]Michael Veale and Lilian Edwards, "Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling" (2018) 34 Computer Law & Security Review 398, 399.

[6]European Commission, *Commission Communication on the Protection of Individuals in Relation to the Processing of Personal Data in the Community and Information Security* (COM (90) 314 final) .

decision making process.[7] This thought of ensuring human participation has been described as an attempt to maintain human dignity, since, instead of an individual data profile, it should be the affected human that influences the outcome of the decision making process.[8] The EC voiced another concern in its 1992 Amended DPD Proposal[9], where it claimed that ADM processes would weaken the sense of responsibility of a human decision maker when interacting with an ADM system. The EC was in the opinion that the "apparently objective and incontrovertible character"[10] of an outcome of an ADM process could lead to non-critical acceptance by the human of the outcome, and little to no incentive to investigate such (potentially flawed) outcomes.[11]

Against this backdrop, art. 15 DPD was adopted in the following form:

<div align="center">Automated Individual Decisions</div>

1. Member States shall grant the right to every person not to be subject to a decision which produces legal effects concerning him or significantly affects him and which is based solely on automated processing of data intended to evaluate certain personal aspects relating to him, such as his performance at work, creditworthiness, reliability, conduct, etc.

2. Subject to the other Articles of this Directive, Member States shall provide that a person may be subjected to a decision of the kind referred to in paragraph 1 if that decision:

   (a) is taken in the course of the entering into or performance of a contract, provided the request for the entering into or the performance of the contract, lodged by the data subject, has been satisfied or that there are suitable measures to safeguard his legitimate interests, such as arrangements allowing him to put his point of view; or

   (b) authorised by a law which also lays down measures to safeguard the data subject's legitimate interests.

2.1.2   (In)Significance of Article 15 DPD

Due to the legislative nature of a directive, the DPD required transposition into national law. Therefore, as opposed to the direct effects of art. 22 GDPR in the EU, art. 15 DPD was implemented differently across EU Member State legislation. Notwithstanding its novelty, the article has played a minor role in the pre-GDPR EU data protection law framework.[12] It has not been a major subject matter in

---

[7]Ibid 29.

[8]Mendoza and Bygrave (n 4) 7.

[9]European Commission, *Amended Proposal for a Council Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data* (COM (92) 422 final, 1992) .

[10]Ibid 26.

[11]Ibid 26.

[12]Veale and Edwards (n 5) 399.

enforcement actions by national data protection authorities, nor has it been addressed for the purposes of adequacy assessments of third countries' data protection regimes.[13] Several reasons led to the practical insignificance of art. 15 DPD. The application of the article required that several conditions are fulfilled; a decision had to be made, this decision had to have legal or other significant effects, the decision had to be based solely on automated processing of personal data, and the purpose of the processing had to be the evaluation of personal aspects of the data subject. If any of these conditions were not met, art. 15 DPD did not apply. Furthermore, the wording of the provision was ambiguous and remained unclear. No official explanatory documents provided guidance, nor did the article become subject of litigation before the Court of Justice of the European Union (CJEU). In 2014, however, the German Federal Court of Justice ("Deutscher Bundesgerichtshof") handed down an appeal judgement in the so-called SCHUFA case[14], which revolved around the use of automated credit-scoring systems. The German court held that such a system did not fall under the national transposition of art. 15(1) DPD. That was because the automated elements of the decision process only served the purpose of preparation of evidence, whereas the decision regarding the grant or denial of a credit was made by a human.[15] Lastly, art. 15(2) DPD provided many exceptions from the rule laid down in paragraph 1, which allowed for an effortless circumvention of the article.[16]

2.2   Rationale of Article 22 GDPR

Notwithstanding the rather minor role and little practical relevance of art. 15 DPD, the essential portion of the article was re-introduced with art. 22 GDPR. The preparatory documents to the GDPR do not contain much information as to the rationale of art. 22 GDPR. However, due to evident similarities to art. 15 DPD, it is reasonable to assume that art. 22 GDPR is a result of similar concerns regarding automation as discussed in Section 2.1.1.[17]

One goal of art. 22 GDPR is to address privacy-related risks emanating from ADM processes. Simultaneously, the legislator did not want to stifle technological development.[18] Furthermore, art. 22(1) GDPR could be linked to considerations of human dignity. That is, when faced with a decision that has legal effects or similarly significantly affects a human, this human should be treated with the dignity of having another human, not a computer, to decide matter.[19] Furthermore, recital 71 to the GDPR addresses concerns regarding the potential of flawed outcomes of ADM processes, and requires that the data controller "ensures in particular, that factors

---

[13]Mendoza and Bygrave (n 4) 4.

[14]*Judgement of the German Federal Court* VI ZR 156/13.

[15]Ibid 34.

[16]Mendoza and Bygrave (n 4) 5.

[17]Ibid 7.

[18]Boris P Paal and Daniel A Pauly, *Datenschutz-Grundverordnung Bundesdatenschutzgesetz* (CH Beck 2017).

[19]Tal Zarsky, "Transparent Predictions" (2013) 2013 University of Illinois Law Review 1503, 1551.

which result in inaccuracies in personal data are corrected and the risk of errors is minimised".

2.3   Analysis of Key Terms

According to recital 15 to the GDPR, the GDPR is supposed to be technology-neutral, meaning that its rules and the protection it provides should not be triggered by the use of a particular technology, but instead set up principles that are applicable for future technological developments.[20] This leads to sometimes vague formulation of the provisions. The upside to this approach is the flexibility and adaptability of the law, while the downside lies in uncertainty regarding the meaning and implications of the law.[21]

In order to better understand art. 22(1) GDPR, the wording *a decision based solely on automated processing* has to be construed. This part of the sentence of art. 22(1) GDPR consists of three main components:

- Component 1: The outcome (*a decision*).

- Component 2: A process to facilitate the decision (*automated processing*).

- Component 3: The outcome is achieved by only relying on the process stated under Component 2 (*based solely*).[22]

Art. 22(1) GDPR is titled *Automated Individual Decision Making, Including Profiling*. In various places, the GDPR refers to ADM without clearly defining what this term means.[23] This lack of definition becomes apparent and problematic when taking a closer look at art. 22(1) GDPR.

2.3.1   Component 1: A Decision

When reading the title of art. 22 GDPR in conjunction with art. 22(1) GDPR, certain inconsistencies regarding the role of the decision taker can be observed. A decision can be understood as a specific stance that is taken with regard to a person, with a high likelihood that this stance is being acted upon.[24] The title of art. 22 GDPR suggests that the article captures decisions performed by an artificial entity. The wording, *a decision based solely on automated processing*, however, leaves open whether the decision taker is human or artificial.

---

[20]Centre for Information Policy Leadership, *Implementing and Interpreting the GDPR: Challenges and Opportunities, Towards a Successful and Consistent Implementation of the GDPR* (2016) ⟨https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_gdpr_amsterdam_ workshop_report___2016.oct.06_.pdf⟩ accessed 2 April 2019, 3, 7.

[21]Emily Pehrsson, "The Meaning of the GDPR Article 22" (2018) 31 European Union Law Working Papers, 5.

[22]Ibid 7.

[23]See Art. 22, 13(2)(f), 14(2)(g), 15(1)(h), recital 71

[24]Mendoza and Bygrave (n 4) 10.

The provision laid down in art. 22(1) GDPR and the title of the article are not coordinated, which complicates a common interpretation of the article. If the issues of ambiguity and lack of clarity in relation to art. 15 DPD are not to be repeated with art. 22 GDPR, such inconsistencies have to be addressed and mitigated. Therefore, it has to be established whether the article refers to a human or an artificial decision maker.

The interpretation that this thesis follows will be provided in Section 2.3.2. That is because the reasoning necessary to assess the entity of the decision taker depends on component 2.

2.3.2   Component 2: Automated Processing

A literal reading of art. 22(1) GDPR produces further contradictions regarding the actual scope of meaning of component 2. A literal reading of the wording *a decision based solely on automated processing* provides the following two possible interpretations of art. 22(1) GDPR:

- The act of decision taking by an artificial entity is considered to be a processing activity. This means that art. 22(1) GDPR prohibits that both the preparatory stage of, and the actual act of taking a decision that is based on the preparation stage, are fully automated. Some authors seems to support this view, stating that "[t]he GDPR [...] targets decisions made through solely automated processing",[25] or "[t]he decisions must result from a process that includes only automated processing, without human intervention."[26] If the act of decision taking is considered to be a form of automated processing activity, the decision making entity would be an artificial one. That is because automated processing pertains to processing that is not conducted by humans, but by automated means.

- Art. 22(1) GDPR prohibits that the act of decision taking is conducted by consulting preparatory outcome generated by automated processing only. The act of decision taking is not considered a processing activity. This interpretation leaves open the question of decision making entity.

It is the author's view that the first interpretation does not apply because it unduly extends the scope of how automated processing is understood in the GDPR. Art. 4(2) GDPR defines processing as "any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction". Automated processing is not explicitly defined in the GDPR, but is mentioned in art. 4(2) GDPR as a processing activity that opposes manual processing

---

[25]Michèle Finck, "Smart Contracts as a Form of Solely Automated Processing Under the GDPR" (2019) 19-01 Max Planck Institute for Innovation & Competition Research Paper Series, 5.

[26]Pehrsson (n 21) 5.

of personal data.[27] A filing software that saves all e-mail addresses of customers that place orders with a company in a database is processing personal data automatically, as no human is manually entering this personal data into a filing system. While art. 22(1) GDPR provides for the right not to be subject to a decision that is based solely on automated processing, it does not provide the right not to be subject to automated processing of personal data in general (automated processing of personal data is, however, governed by special data protection rules, such as art. 20(1)(b) and art. 35(3)(a) GDPR).

The author argues that the act of decision taking does not constitute a processing activity, i.e., automated processing of personal data and the act of decision taking are two separate acts. The definition of *processing* in art. 4(4) GDPR pertains to tasks such as collecting, organising, and storage of personal data. These tasks do not require high level evaluation of facts in order to make impactful decisions such as whether a person is to be considered creditworthy. The act of organising personal data might contain some form of decision taking, e.g., observing and subsequent categorisation of a data subject due to age and income. However, this is an observation of given facts, not a judgement whether or not the person ultimately is deemed creditworthy. If the act of decision taking were to fall under the GDPR's processing definition, the definition of art. 4(2) GDPR would be broadened significantly.

The second interpretation does not provide a reasonable explanation of art. 22(1) GDPR either. Automated processing produces an output. These outputs then translate with or without human participation to decisions.[28] That means that automated processing merely serves a preparatory function. In order to argue that the prohibition of fully automated preparatory stages is not what the EU legislator had in mind, reference is made to the SCHUFA case[29] pointed out in Section 2.1.2, where the German Federal Court of Justice held that due to the fact that automated processing was only used for preparation of evidence, the credit-scoring system was not an ADM process.[30]

While art. 22(1) GDPR literally refers to *a decision based solely on automated processing*, this thesis takes the view that art. 22 GDPR in fact aims to capture the act of decision taking which is conducted by an artificial entity. As shown above, any other interpretation would lead to contradictory outcomes. Therefore, the wording *a decision based solely on automated processing* should be interpreted as to mean *a solely automated decision*. Therefore, this thesis uses the latter term from this point on.

2.3.3 Component 3: Based Solely

The previous section argues that *a decision based solely on automated processing* should be understood as *a solely automated decision*. The term *solely* is a decisive part of the latter wording, since it provides a statement regarding the degree of

---

[27]See recital 15 to the GDPR; art. 4(2) GDPR

[28]Veale and Edwards (n 5) 2.

[29]*SCHUFA case* (n 14).

[30]Ibid.

automation in the decision making process. In order to distinguish solely and not solely automated decision for the purposes of art. 22(1) GDPR, the degree of human involvement in the decision making process is crucial.

Two understandings of *solely* are possible. A narrow understanding of *solely automated decision making* requires absence of human participation in the decision making process. The implication of this strict approach to *solely* is that any type of human conduct, as insignificant as it may be, renders a decision not solely automated, and thereby prevents the application of art. 22 GDPR. The narrow approach means that the prohibition contained in art. 22(1) GDPR can be easily circumvented by inserting nominal human conduct into the decision making process.[31] Therefore, if the goal of the legislator is to promote the enforcement of art. 22 GDPR, a narrow understanding of *solely* should be rejected.

Higher requirements are set up by a broad understanding of *solely automated decision making.* A broad understanding requires that the degree of involvement of a human meets a certain threshold in order to render a decision not solely automated. As opposed to the narrow understanding, not any kind of human conduct suffices. The implications of the broad approach to interpreting *solely* are that art. 22(1) GDPR remains applicable in cases of nominal human conduct, and that the circumvention of the prohibition is more difficult.

2.4   Qualified Human Participation

A broad understanding of *solely* means that human participation in the ADM process has to meet a threshold in order for this process not to be considered solely automated. Thit means that criteria have to be established which determine whether a human is sufficiently involved.

2.4.1   Recitals to the GDPR

The only recitals that pertain to the interpretation of art. 22 GDPR are recital 70, 71, and 72. However, only recital 71 contains a statement relating to qualified human participation: "...[t]he data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing and which produces legal effects concerning him or her or similarly significantly affects him or her, such as automatic refusal of an online credit application or e-recruiting practices *without any human intervention* (emphasis added)." These statements constitute a mere paraphrasing of the wording in art. 22(1) GDPR, since *solely automated* and *without any human intervention* ultimately mean the same, that is, the decision making process is fully automated. Therefore, the GDPR does not provide any help to develop a definition of qualified human participation.

---

[31]Sandra Wachter and others, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" (2017) 7 International Data Privacy Law 76, 88.

2.4.2   Official Guidance

Up to this day, only two official authorities have attempted to explain the term *a solely automated decision*. The first is the EDPB in its Guidelines on Automated Individual Decision Making and Profiling for the Purposes of Regulation 2016/679 from February 6, 2018.[32] The second is the British national supervisory authority, ICO, which provides information on this matter on its website.[33]

European Data Protection Board

In the opinion of the EDPB, *solely* means that there is "[...] no human involvement in the decision process."[34] The EDPB further states that is not possible to avoid the prohibition of art. 22(1) GDPR by "fabricating human involvement"[35]. The following example is provided to explain fabricated human involvement:

"(...) (I)f someone routinely applied automatically generated profiles to individuals without any actual influence on the result, this would still be a decision based solely on automated processing."[36]

Human involvement further requires that "any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision."[37]

While these statements by the EDPB provide for a general direction in which an understanding of qualified human participation could go, the shortcomings of these statements are numerous.

The first one is prioritisation. In its Guidelines on Automated Individual Decision Making and Profiling for the Purposes of Regulation 2016/679, which consists of 37 pages, not even half a page is dedicated towards the meaning of *solely*. Instead, the EDPB engages in long elaborations on the explanation of the term *legal or similarly significant effects* contained in art. 22(1) GDPR, or on the exceptions of the prohibition contained in art. 22(1) GDPR. While these issues are important for the correct application of art. 22 GDPR, it is surprising that the EDPB does not elaborate more thoroughly on the term *solely*. As pointed out in Section 2.1.2, two related reasons for the little practical application of art. 15(1) DPD were the lack of an EU-wide official guidance document, and the ambiguous formulation of the article. In order to avoid a similar scenario regarding art. 22 GDPR, relevant authorities such as the EDPB should step in and provide definitions. However, by choosing

---

[32]Article 29 Working Party, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* (WP251, 2017) .

[33]"What does the GDPR say about automated decision-making and profiling?" (*Information Commissioner's Office*) ⟨https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/automated-decision-making-and-profiling/what-does-the-gdpr-say-about-automated-decision-making-and-profiling/⟩ accessed 20 March 2019.

[34]*WP251* (n 32) 20.

[35]Ibid 21.

[36]Ibid 21.

[37]Ibid 21.

to not delve deeper into the explanation of the term *solely*, the EDPB missed an opportunity to clarify art. 22(1) GDPR.

The second shortcoming is the lack of helpful clarification. By demanding "actual influence on the result"[38], meaningful oversight over the decision, and performance of such oversight by a person vested with "the authority and competence to change the decision"[39], the EDPB provides for a few starting points of discussion. However, the EDPB fails to provide more detailed explanation about what these terms mean. For example, when referring to the requirement of meaningful oversight, the EDPB states that the person exercising the oversight should have the "authority and competence to change the decision"[40]. Depending on the organisational structure of an undertaking, a person vested with authority to change the decision might be a person on a very high hierarchy level. This would render the requirement unfeasible, since this person would most probably not be able to supervise each and every single decision that is being made. A useful addition would have been if the EDPB defined the sphere of humans it considered to be authorised to change the decision. Furthermore, competence can be interpreted differently, e.g., as referring to specialised knowledge, work experience, or mental capacity. It is also unclear how competence would have to be evaluated. Lastly, the EDPB refrains from pointing out that its explanation require further development, which could have motivated further research on the term *solely*.

While certainly difficult to accomplish in a section that contains less than 200 words, the EDPB nevertheless fails to establish clear requirements for the required influence, oversight, authority, and competence. Instead of an explanation, more notions that require interpretation are produced, and the EDPB has missed one important opportunity to set up a threshold that allows one to grasp the exact meaning of human intervention.

Information Commissioner's Office

Similar to the EDPB, the ICO states that the term *solely* means that "a decision-making process [...] is totally automated and excludes any human influence on the outcome."[41] The ICO clarifies that influence can mean that "someone weighs up and interprets the results of an automated decision before applying it to the individual."[42] Furthermore, human involvement must be "active, meaning that a human should review the decision before it is applied and has discretion to alter it."[43]

Similar to the EDPB, the extent to which the ICO elaborates on the term *solely* is limited. However, the ICO provides more details and somewhat concrete

---

[38]*WP251* (n 32) 21.

[39]Ibid 21.

[40]Ibid 21.

[41]"Rights related to automated decision making including profiling" (*Information Commissioner's Office*) ⟨https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/⟩ accessed 20 March 2019.

[42]"What does the GDPR say about automated decision-making and profiling?" (n 33).

[43]Ibid.

suggestions on how human involvement should look like, e.g., someone weighing up and interpreting the result of a decision. Who such person should be, however, remains unclear.

2.5 Chapter Conclusion

Art. 22(1) GDPR has attracted much academic attention, is subject to official non-binding explanatory guidance, and may be read in the light of earlier sources in relation to art. 15 DPD. However, ambiguities regarding art. 22 GDPR remain.

The wording, *a decision based solely on automated processing*, contradicts the rationale of the article. For this reason, this thesis suggests to use the wording, *a solely automated decision*. Furthermore, the meaning of the term *solely*, i.e., the threshold for qualified human participation, is unclear. Such ambiguities will prevent art. 22(1) GDPR from having any more impact than its predecessor. Therefore, the lack of clarity has to be resolved.

Chapter 3

Defining Qualified Human Participation

This chapter investigates how MHC as defined by Filippo Santoni de Sio and Jeroen van den Hoven[1], and MHA as defined by Ben Wagner[2], aid to define qualified human participation for the purpose of art. 22(1) GDPR. Each set of requirements is analysed, the concepts are compared, and their strengths and weaknesses are identified. In order to overcome the respective weaknesses of the two concepts and adapt them to a GDPR setting, a new approach is introduced, called Extended MHC, which combines aspects of both MHC and MHA.

3.1   Meaningful Human Control

MHC is a concept that emerged from the debate on the deployment of fully autonomous weapon systems (AWS) in armed conflicts.[3] There is currently no internationally agreed definition of AWS, but they are often described as machines that, after initial activation by a human operator, are able to select and engage targets without further participation by the operator.[4] The prospect of AWS has caused societal concern regarding the current and future ability of the technology to conform with International Humanitarian Law principles,[5] whether or not the deployment of such technology is ethical, and potential issues regarding the attribution of legal

---

[1]Filippo Santoni de Sio and Jeroen van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account" (2018) 5 Frontiers in Robotics and AI 1.

[2]Ben Wagner, "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems" (2019) 11 Policy & Internet 104.

[3]"Killer robots: UK Government Policy on Fully Autonomous Weapons" (*Article 36*, 2013) ⟨http://www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf⟩ accessed 20 May 2019, 1.

[4]United States Department of Defense, *Department of Defense Directive 3000.09: Autonomy in Weapon Systems, November 21, 2012* ⟨https://www.hsdl.org/?abstract&did=726163⟩ accessed 5 April 2019.

[5]Thompson Chengeta, "Defining the Emerging Notion of "Meaningful Human Control" in Weapon Systems" (2016) 49 NYU Journal of International Law 833, 835.

liability.[6] Opposing views stress the desirability of the deployment of AWS due to its potential to reduce causalities on the battle field.[7]

The proposed ways to handle the prospective deployment of AWS differ. While some advocate for a preemptive ban of AWS by means of an international treaty,[8] in 2013 the British NGO Article 36 introduced the now prominent concept of MHC. The NGO claims that MHC over the use of lethal force must be maintained. This means that humans, and not artificial systems, should ultimately remain in control of, and thus be morally responsible for decisions about lethal military operations.[9] MHC was a much debated topic of discussion during sessions of the Group of Governmental Experts (GGE) on emerging technologies in the area of Lethal Autonomous Weapons Systems (LAWS) of the High Contracting Parties to the Convention on Prohibitions or Restrictions on the Use of Certain Conventional Weapons (CCW).[10] States that are party to the CCW broadly agree that retaining human control over AWS is paramount.[11] While various authors have set forth different explanations of MHC, there is no internationally agreed definition of MHC.

3.1.1   MHC by Santoni De Sio and Van Den Hoven

The first approach to define qualified human participation is one specific account of MHC. In the paper "Meaningful Human Control Over Autonomous Systems: A Philosophical Account", Filippo Santoni de Sio and Jeroen van den Hoven define two requirements for MHC. The authors consider MHC as a crucial notion in ethics discussion on AI and robotics, and advocate for an application of this concept over autonomous robotic systems in general.[12]

The reason for the choice of this specific account is twofold. First, the authors have set up two conditions in order for human control to be meaningful. This enables a structured approach to adopting the concept to a GDPR setting. Second, the authors claim that their account of MHC can be applied beyond AWS, hence broadening the scope of application of MHC. It is therefore valuable to assess whether Santoni de Sio's and van den Hoven's account can resolve the issue stated in Chapter 2. That is, to define qualified human participation, and thereby establish legal certainty regarding the meaning of *a solely automated decision* for the purposes of art. 22(1) GDPR. In the following sections, references to MHC point out Santoni de Sio's and van den Hoven's specific account of MHC, not the idea of MHC in general.

---

[6]Santoni de Sio and van den Hoven (n 1) 2.

[7]Ronald C Arkin, "Lethal Autonomous Systems and the Plight of the Non-Combatant" in Ryan Kiggings (ed), *The Political Economy of Robots, Prospects for Prosperity and Peace in the Automated 21st Century* (Springer-Cham Palgrave Macmillan 2018) 3.

[8]"Campaign to Stop Killer Robots" ⟨https://www.stopkillerrobots.org/⟩ accessed 18 May 2019.

[9]"Killing by Machine: Key Issues for Understanding Meaningful Human Control" (*Article 36*, 2015) ⟨http://www.article36.org/wp-content/uploads/2013/06/KILLING_BY_MACHINE_6.4. 15.pdf⟩ accessed 20 May 2019.

[10]*Report of the 2018 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems* ⟨https://undocs.org/en/CCW/GGE.1/2018/3⟩ accessed 29 January 2019.

[11]Ibid.

[12]Santoni de Sio and van den Hoven (n 1) 11.

3.1.2   Two Conditions for MHC

Santoni de Sio and van den Hoven suggest that the following two conditions should be fulfilled in order for an ADM system to be considered to be under MHC:

1. Tracking condition: An ADM system must be programmed in a way that enables the system to be responsive and adaptable to factual, legal, and moral considerations of a human that designs or operates the system. This means that an ADM system must be able to observe relevant human reasons in a given situation, and manipulate its behaviour according to the observed considerations. For example, when applied to an AWS setting, the tracking condition requires that the system tracks, inter alia, considerations flowing from international law such as necessity, proportionality and discrimination, and is able to adapt its behaviour according to morally relevant considerations such as the distinction between civilian and civilian involved in an armed attack.

2. Tracing condition: The behaviour of an ADM system must be traceable to technical and moral comprehension of at least one human that either designs or operates the system. This human must comprehend or be able to comprehend the capabilities and the possible effects of the use of the system. Furthermore, the human must comprehend or be able to comprehend that the use of the system or its effects on the world lead to reactions from other humans. This means that the human agent must be aware of its responsibility for the actions of the ADM system.[13] The tracing condition applied to the AWS requires that the military staff has a proper understanding of the capability, the functioning, and the limitations of the AWS.[14]

3.1.3   Analysis of MHC Conditions

In the following sections, the tracking and tracing conditions are analysed, and strengths and weaknesses for the purpose of defining qualified human participation are identified.

Tracking Condition

Section 2.4.2 discusses the notion of influence proposed by the ICO. A strength of the ICO's guidance on explaining *a solely automated decision* is that the ICO not only suggests a condition for qualified human participation, but also explains how to fulfil the condition. The aim of the tracking condition is alignment of the system's behaviour with human considerations that are relevant in a given case. The way the system achieves this goal is by adapting its behaviour to such considerations. The conditions of MHC are formulated as a design requirement, i.e., the tracking and tracing condition have to be programmed for and incorporated into the system's

---

[13]Ibid 6-8.

[14]Dan Saxon, "Autonomous drones and individual criminal responsibility" in Ezio Di Nucci and Filippo Santoni de Sio (eds), *Drones and Responsibility: Legal, Philosophical, and Socio-Technical Perspectives on the Use of Remotely Controlled Weapons* (Routledge 2016).

code.[15] The fact that both the aim and the approach to achieve this aim are stated by the tracking condition makes it easier to grasp and put the tracking condition into practice.

An important aspect of the tracing condition is that it does not require that the ADM system has a capacity to reason.[16] The entity that engages in reasoning remains the human designer or operator, and the system's task is to adapt its behaviour to such human reasoning. This establishes a reasonable expectation regarding the design of automated systems and the control over such systems, since programming an ability to align with human considerations seems more feasible than creating a system that is able to engage in reasoning itself. The tracking condition therefore sets up a realistic and short-term goal that bears the potential to motivate research and further elaboration on the issue of a tracking ability.

MHC is highly context-dependent, i.e., what constitutes relevant considerations, the level of sufficient responsiveness to these considerations, and the human agent that the system should adapt to cannot be established generally.[17] Instead, a case-by-case assessment is necessary.

One approach to identify relevant considerations is to first define the task that the ADM process should fulfil. Once the task is identified, the considerations necessary to fulfil the task in a desirable manner can be established. This in turn gives an indication as to which humans are suitable for the role of the agent. As long as the agent is human, relevant agents can be the programmer or operator of the system, but also legislators or policy-makers who govern the system.[18] This results in a wide range of possible agents, starting with the programmer, a human that develops the system, and going as far as a human that sets up rules or guidelines regarding the use of the system. Lastly, Santoni de Sio and van den Hoven point out that it may be necessary to establish a threshold for evaluating how much reliability in the system is necessary, i.e., to define how well the system has to be able to respond to the relevant considerations in order to fulfil the tracking condition.[19] This means that the application of the tracking condition requires a decision regarding the quality and effectiveness of the responsiveness of an ADM system.

Due to the various possible GDPR settings in which ADM can be deployed, this context-dependent approach is important. For example, an ADM process that engages in credit-scoring has to track different human considerations than an ADM that is deployed in a employment process. While the former takes into consideration circumstances such as income and debts, the latter takes into account work experience and educational background.

As a result, the tracking condition provides the following roadmap to establish qualified human participation:

1. Identify the relevant considerations to be tracked by the system.

---

[15]Santoni de Sio and van den Hoven (n 1) 3.

[16]Ibid 7.

[17]Ibid 11.

[18]Ibid 8.

[19]Ibid 7.

2. Define which human agents the system should track.

3. Establish the amount of responsiveness of the system.[20]

The tracking condition clarifies the notion of meaningful oversight of the EDPB that was discussed in Section 2.4.2. The EDPB states that in order for human conduct to be considered sufficient, it is necessary to ensure that "any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision."[21] In order to fulfil the tracking condition, the ADM process must align its conduct to relevant human considerations. If the system does not engage in tracking, it cannot be considered to be under MHC. This dependency of the system on the human agent, and the requirement that the system adapts to the agents, can be seen as one possible form of oversight. This oversight is meaningful, since both the absence of a tracking ability or of the agent lead to the outcome that MHC is not fulfilled.

Furthermore, the EDPB requires that the person that is exercising meaningful oversight must be vested with authority and competence to change the decision.[22] The tracking condition requires the system to follow the reasoning of the human. If, for example, the human agent wrongfully assumes that the subject of a credit-scoring process is a person of age who has no regular income, but in fact is a minor that is not engaging in any paid work, the tracking condition accommodates for such a change of relevant considerations by adapting to the agent's reasoning, which would be that the subject should not be scored. If the relevant considerations of the agent change, the tracking condition makes sure that the ADM process follows up, which means that the possibility to change a decision is given.

Tracing Condition

The tracing condition requires comprehension of the capabilities and possible effects of the use of the system, and awareness that the use of the system and its effects on the world lead to reactions from other humans.[23] This means that the group of humans capable to fulfil the tracing condition is restricted to humans who have a proper understanding of both the moral considerations regarding the deployment of an ADM process, and also the technical capabilities of the system. As a consequence, a highly talented computer scientist who does not take relevant moral reasoning into consideration does not meet the threshold. Neither does an ethicist who does not understand the functioning and effects of the system.

Section 2.4.2 points out that the EDPB does not clarify what its proposed notions of authority and competence for the purpose of meaningful human oversight mean. The tracing condition resolves this omission, since it defines the group of humans who are authorised and competent to exercise control. Furthermore, the tracing

---

[20]Ibid 11.

[21]Article 29 Working Party, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* (WP251, 2017) 21.

[22]Ibid 21.

[23]Santoni de Sio and van den Hoven (n 1) 8.

condition addresses concerns of the European legislator regarding automation that
are discussed in Section 2.1.1. The EC anticipated that ADM weakens the sense of
responsibility of a human that interacts with an ADM process.[24] Particularly, the
EC was afraid that the human would attach too much weight to the outcome of an
ADM process, and therefore would not critically assess such an outcome.[25]

The tracing condition mitigates this concern, since it requires moral awareness
and technical competence of the staff that is designing or operating an ADM system.
If the human operator is aware of the fact that the ADM system's behaviour traces
back, the human operator is most likely to be more engaged in participating critically
in the decision making process, and has a strong incentive to investigate any errors
produced by the ADM system. The tracing condition means that companies that
deploy ADM systems have to provide relevant moral and technical training to
promote its personnel's understanding of risks and responsibilities associated with
programming and deploying ADM processes.[26] Therefore, the tracing condition may
in fact lead to an increase of awareness of responsibility for humans involved in
human-machine interactions.

3.2   Meaningful Human Agency

The second approach to define qualified human participation is MHA. Wagner argues
that effective protection of human rights requires that humans remain an integral part
of ADM systems.[27] For these purposes, Wagner suggests that in cases in which the
human is only included in an ADM process as a "basic rubber-stamping mechanism"[28],
such processes should be considered quasi-automated.[29] Quasi-automation means
that a human is not able to exercise any actual influence on the result of an ADM
system, i.e., that the human has no agency.[30]

3.2.1   Conditions for MHA

Wagner defines seven criteria to be taken into account when assessing if an ADM
system is under MHA:

  1. Time element: This criteria relates to the amount of time that the human has
     in order to conduct a certain task. The less time the agent has, the more likely
     it is that the process is quasi-automated.

---

[24]European Commission, *Commission Communication on the Protection of Individuals in Relation
to the Processing of Personal Data in the Community and Information Security* (COM (90) 314
final) 29.

[25]Isak Mendoza and Lee A Bygrave, *The Right Not to Be Subject to Automated Decisions Based
on Profiling* (Springer 2017) 5.

[26]Santoni de Sio and van den Hoven (n 1) 12.

[27]Wagner (n 2) 108.

[28]Ibid 1.

[29]Ibid 114.

[30]Ibid 11.

2. Qualification element: This criteria addresses the level of qualification of the human. The less qualified the agent is to perform the task, the more likely it is that the process is quasi-automated.

3. Liability element: This criteria pertains to the distribution of liability. Wagner argues that the greater the amount of liability assigned to the human agent in case of a failure of the system, the more likely it is that a process is quasi-automated. The liability element is based on the assumption that a large amount of legal liability assigned to the human is an indication that the human has only been introduced into the process in order to attract potential legal consequences in case of an error of the system.

4. Support element: This element refers to the amount of external help the human agent can rely on when fulfilling a task. Some tasks require a high degree of concentration and effort, or lead to negative effects on the agent's psyche. External help should support the human when making a decision, and prevent negative effects on the agent. The higher the level of support, the less likely it is that a process is quasi-automated. External support refers to, for example, psychological support after traumatising decisions.

5. Adaption element: The more the human agent has to adapt to the system, and the less the system is designed around the human agent, the more likely it is that a system is quasi-automated.

6. Information element: The information element requires that the human agent has the possibility to access all relevant information in order to fulfil a task correctly. If this is not the case, a process is likely to be quasi-automated.

7. Agency element: The agency element requires that the human agent must have the authority to change the decision, and exercises this authority regularly. The weaker the authority to change a decision, and the less the human agent exercises this authority, the more likely it is that a process is quasi-automated. If the only function of the human operator is to regularly agree with the machine and only very rarely disagrees with it, it is highly likely that the human operator's agency is insufficient.[31]

3.2.2 Analysis of MHA Conditions

In the following sections, the seven MHA elements are analysed, and strengths and weaknesses for defining qualified human participation are identified.

Time Element

The time element is based on the assumption that a human agent performs better when provided with sufficient time to conduct a task. The need to take immediate actions can lead to stressful and overwhelming situations for the human. Depending

---

[31]Ibid 115.

on the agent's ability to deal with pressure, such a situation can render the agent's conduct flawed due to sloppiness or inaccuracy. Another consequence of lack of time can be that the human agent is more likely to hand over some parts of, or even the entire task, to the ADM system in order to save time or to be able to complete a task.[32] If the human agent delegates the entire task, any human participation in an ADM process would be eliminated, which would render the process fully automated. This omission would be contrary to the purposes of art. 22(1) GDPR.

The time element constitutes a clear condition regarding the explanation of qualified human participation. To ensure that the human has enough time to conduct a task can prevent flawed outcome and delegation of tasks due to time constraints. If the human is under time pressure when conducting a task, human participation is not considered to be qualified for the purposes of art. 22(1) GDPR. However, the exact time frames are not defined by the time element. It has to be taken into account that some tasks may require very little time to be conducted. An adaptation of the time element that accounts for such low-effort tasks would be necessary. Another challenge is that the assessment of how much time is necessary to conduct a task is highly subjective. That means that opinions regarding the extent of time may differ from person to person. Therefore, time frames for types of tasks have to be defined in order to adapt the time element to a GDPR setting. Other options are to evaluate the necessary time individually in an exchange with the human, which allows the possibility to adapt the time frame to personal traits of the human.

Qualification Element

The qualification element ensures that human agents are equipped with the required competences to fulfil a task. The qualification element is in line with the EDPB's notion of competence, and the tracing condition.

The critique mentioned in Section 2.4.2 regarding the EDPB's notion of competence is valid for the qualification element, too. That is, the type of competence or qualification necessary are unclear. A lack of qualification can take various forms. For example, the human agent can lack the necessary knowledge and expertise that is required to deal with a certain task. Another possibility is that a human agent that is new at a certain profession does meet the formal qualification requirements, but lacks crucial work experience. Yet another possibility is that the agent does not have the personality traits that are necessary to be able to deal with a certain task, such as lack of patience or bad temper.

In order to apply the qualification element in a useful way, the relevant qualifications for each ADM process have to be defined. Similar to what the tracing condition requires, qualification can require training to improve the human's moral and technical awareness of the consequences of the deployment of an ADM system. Then, in order to ensure that the qualification requirements are being followed, a screening process has to be put in place which assesses whether the human meets the requirements.

---

[32]Linda J Skitka and others, "Accountability and Automation Bias" (2000) 52 International Journal of Human-Computer Studies 701.

Liability Element

The sources that were analysed in Chapter 2 do not mention a criterion that is similar to the liability element. Wagner proposes that a large amount of liability indicates that the human has only been inserted into the ADM process to attract responsibility for errors of the system, i.e., that the role of the human is restricted to attracting blame. Therefore, this MHA requirement would constitute a novel criterion to explain qualified human participation.

Wagner fails to explain what failure of the system means. Since liability is triggered with such failure, it can be assumed that the implications of a failure have to amount of serious consequences. Reference can be made to the EDPB's explanation of legal and similar significant effects: in order for a solely automated decision to fall under the scope of art. 22 GDPR, the decision must produce either legal effects that concern the data subject, or similarly significantly affect the data subject. While the GDPR does provide a definition of legal effects, the EDPB assumes that only serious, impactful effects are referred to, i.e., effects that pertain to the data subject's legal rights or its legal status. Examples of such legal rights are cancellation of a contract, the entitlement to or denial of a social benefit granted by law, or refusal of admission to a country or denial of citizenship.[33]

It is questionable whether attributing legal liability to a human that is not serving any other purpose than attracting liability would hold in front of a court. If a judge has clear evidence that the human has not been responsible for the error of the ADM system, and is only involved to attract any adverse legal consequences, it is highly unlikely that the human would in fact be held liable for the error. Furthermore, Wagner does not define whether the liability element covers criminal and/or civil liability.

Support Element

The possibility to rely on external support is one way to ensure that the human has the opportunity to participate in an active and meaningful manner in the ADM process. Wagner does not provide a final list of possible forms of support. Support can take many different forms, for example as interaction in the form of discussions with co-workers with similar tasks, feedback from supervisors, or the possibility to consult psychological help after traumatising tasks. In order to assess whether the human has sufficient possibility to rely on external support, more detailed elaborations on what forms of support are necessary for each particular human-machine interaction have to be made. Furthermore, it has to be established at what stage during the fulfilment of the task such support has to be offered. Furthermore, the type of support and the duration of such support has to be defined. However, the subjectivity of what constitutes necessary and timely support makes it difficult to establish any general rules. Therefore, a possible approach to find out what form of support is necessary in a given human-machine interaction is to consult with the human agents that

---

[33] *WP251* (n 21) 21.

participate in the process. Such a case-by-case analysis provides for personalised support mechanisms.

Adaption Element

This criterion differs from the other six elements of MHA, since it formulates a rule regarding the architecture of the ADM process. The adaption element requires that the ADM process should be designed in a way that the automated system adapts to the human, rather than the other way around. In relation to the time element, for example, this can mean that the ADM process has to programmed in a way that enables the system to accommodate for the necessary amount of time required by the human to conduct a task.

The adaption element has a similarity with the MHC approach, where the conditions are formulated in the form of design requirements. Similarly, the adaption element means that the adaption of the ADM has to be incorporated into the system's code. This element of MHA is a good example of how undertakings would be required to change their technological processes in order to adhere to art. 22(1) GDPR, that is, to change how ADM processes are designed.

The weakness of the adaption element is that Wagner does not provide any solution as to how such adaption can be ensured. While the goal of the adaption element seems reasonable, it is a very abstract objective which would require further clarification regarding the means to achieve it.

Information Element

The information element follows the EDPB's statement that persons with the authority and competence to change a decision must consider all relevant data when changing a solely automated decision.[34] The information element pertains to the necessity of the human to be able to assess a given situation correctly in order to fulfil a task. If the human is ascribed with a decision, but not able to actively assess the information in order to formalise a decision, the human lacks the necessary background knowledge.

One big challenge in relation to the information element is to establish what *relevant* information means. Wagner does not engage in any further elaboration. Opinions as to what constitutes relevant information may differ depending on the human, i.e., that the scope of relevance has to be assessed on a case-by-case basis. Therefore, the information element does not provide any further clarification for explaining the threshold between a solely or not solely automated decision. Instead, it confirms that the human involved in an ADM process has to have access to a sufficient amount of information in order to exercise qualified human participation.

Agency Element

The agency element relates to the statement of the EDPB regarding oversight over a solely automated decision, according to which such oversight should be carried

---

[34] *WP251* (n 21) 21.

out by someone who has the authority to change the decision.[35] As Section 2.4.2 points out, the term authority remains undefined by the EDPB. Since the agency requirements is directly referring to the EDPB's statement, it does not provide any further explanation of what authority means. Therefore, the agency requirement does not contribute in any novel or explanatory way in order to define qualified human participation.

Another weakness of the agency element is the requirement that the human exercises the authority to change a decision regularly. This means that the human has to frequently change the decisions that are formalised by the ADM system. By stating this, Wagner suggests that the fulfilment of the agency element depends on that the occasions on which the human intervenes are often. This requirement misses the point of agency. The motivation to change the decision should be to make interventions based on reasonable grounds, not to fulfil a quota of interventions. This becomes apparent in a situation in which an ADM process outperforms the human in correctly performing a given task. If the human is required to frequently disagree, just in order to fulfil the agency requirement, an undesirable situation is created in which human intervention in fact becomes pointless. Therefore, the agency element is not considered suitable to define qualified human participation.

3.3   Comparison of MHC and MHA

The concepts of MHC and MHA emerged from different backgrounds. MHC originates from discussions regarding the deployment of AWS, and was proposed as a reaction towards the ban of such systems,[36] whereas MHA proposes how to integrate human rights into the infrastructure of the internet.[37] According to recital 71 to the GDPR, art. 22(1) GDPR addresses issues such as automatic refusal of an online credit application or automated e-recruiting practices. The GDPR is concerned with the processing of personal data in a civilian setting, and therefore has a similar background as MHA. The military setting of MHC, however, does not mean that the concept is not applicable beyond a this setting.

The GDPR, MHC and MHA pursue the same goal, which is the protection of human rights. Both MHC and MHA pertain to the idea that such protection requires that a human exercises a certain amount of influence over an ADM process. Santoni de Sio and van den Hoven stress that a high degree of system autonomy must be combined with human control, responsibility, and accountability.[38] Wagner argues that human rights in a social-technical system setting can only be protected if this system provides the human with the autonomy to make decisions.[39] MHC tries to protect human rights by establishing responsibility and accountability, while MHA aims to preserve human autonomy.

---

[35]Ibid 21.

[36]Santoni de Sio and van den Hoven (n 1) 11.

[37]Wagner (n 2) 106.

[38]Santoni de Sio and van den Hoven (n 1) 2, 11.

[39]Wagner (n 2) 116.

As pointed out in Section 2.1.1, art. 15 DPD flowed from concerns in relation to automation. Such concerns were that ADM processes would lead to a decline of human influence on shaping important decisions, or a decrease of the sense of responsibility of humans involved in ADM processes.[40] MHC and MHA address and mitigate these concerns. The two concepts convey the thought that mere human presence in an ADM process is not enough to establish sufficient human control over an ADM process. Instead, the degree of human conduct has to be above a certain threshold. If this threshold is met, control is considered to be *meaningful.* The tracing condition, for example, requires moral and technical competence, and awareness of responsibility of the staff that is designing or operating an ADM system. This threshold approach is in line with a broad understanding of the term *solely* for the purposes of art. 22(1) GDPR, according to which only a qualified type of human conduct evades the application of the article. Due to the similarity of the general objectives of MHC and MHA to the aim of art. 22(1) GDPR, and the requirement of a certain degree of human participation in the ADM process, MHC and MHA are useful sources to clarify qualified human participation for the purposes of art. 22(1) GDPR.

While MHC and MHA pursue the same objective, the emphasis of the means in order to achieve this objective vary. MHC aims to establish responsibility and accountability of humans designing or operating ADM systems,[41], while MHA is targeted at preserving human autonomy in and ADM process.[42] In relation to this difference, the different backgrounds of the concepts are relevant. MHC tries to mitigate the problem of the so-called responsibility gap.[43] The term captures the idea that attribution of moral and legal responsibility for crimes or accidents in armed conflict becomes difficult if an autonomous system is deployed. That is due to the fact that traditional legal condition which attract liability might not be fulfilled if an artificial entity is introduced.[44] In order to bridge the responsibility gap, the tracking and tracing conditions aim to maintain a sufficient amount of human control over an ADM system, which in turn help to attribute responsibility for the behaviour of the ADM system. The way in which the two concepts deal with responsibility significantly differs. As opposed to MHC, the focus of MHA is not the idea of ascription of responsibility, but to promote human autonomy in an ADM process. The liability element of MHA assumes that attribution of responsibility to the human is a sign that the process is quasi-automated, i.e., that the human is not exercising significant control over the system.[45] For MHC, successful ascription of responsibility means that human control is established, while for MHA responsibility is an indication that human participation is not meaningful.

Furthermore, the proposed approaches of MHC and MHA vary. The tracking and tracing condition are formulated in the shape of design requirements for automated systems, i.e., not only does a human have to fulfil the two conditions, but the ADM

---

[40]*COM (90) 314 final* (n 24) 29.

[41]Santoni de Sio and van den Hoven (n 1) 2, 11.

[42]Wagner (n 2) 116.

[43]Andreas Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata" (2004) 6 Ethics and Information Technology 175.

[44]Santoni de Sio and van den Hoven (n 1) 4, 5.

[45]Wagner (n 2) 115, 116.

process has to be programmed and designed in a way that enables the human to fulfil the two conditions.[46] This approach follows the notions of Responsible Innovation and Value-sensitive Design, which suggest that ethical considerations regarding a technology should be accounted for at an early stage of the life cycle of the technology, i.e., when it is being developed.[47]

The advantage of the design approach is that at the moment the ADM system is deployed, the possibility for MHC is, at least from the technical perspective, already provided for. As pointed out in Section 3.1.3, the implementation of the tracing condition requires a rethinking and remodelling of current practices of companies that deploy ADM.[48] That is because in order to have the required technical and moral understanding of the ADM process and its implications on its environment, the staff that designs and operates the ADM system has to be highly knowledgeable and educated, which can be ensured by relevant training offered by the employee. Such training would leads to increased awareness of both technical and moral issues among the staff. However, this process is bound to be time consuming, which means that a fast implementation of MHC for the purposes of art. 22(1) GDPR is unlikely. The time it takes for a company to adopt the tracking and tracing condition depends on various factors. Internal factors can be financial and human resources, and willingness to change current practices. External factors can be the current state of knowledge and expertise required to program tracking and tracing, or the amount of regulatory pressure exercised to push the compliance with art. 22(1) GDPR. The latter depends on the significance that art. 22 GDPR will have in the coming years. A long time span until actual implementation of the MHC conditions means continuous uncertainty regarding the effects of the article up to the point of implementation.

The MHA elements can be read as design requirements, too. For example, the time element could mean that the ADM process has to be designed in a way that ensures that the human has sufficient time to conduct a task. However, Wagner does not specifically state that the seven elements have to be incorporated into the architecture of the ADM process. One exception is the adaption element, which sets up the requirement that the system must be designed around the human.[49] The fact that the concept of MHA does not provide any general suggestion regarding when and how the elements should be implemented is a clear disadvantage compared to MHC, since it remains somewhat unclear at what stage of the deployment of an ADM process the seven MHA elements have to be implemented.

The fact that MHA is not explicitly intended to be incorporated into the design of an ADM process, however, also constitutes an advantage over MHC. The tracking condition of MHC requires that the ADM system is capable of tracking relevant human reasoning. That means that the ADM process has to have an alignment ability. This ability first has to be designed, tested, and ultimately implemented into the ADM process. The elements of MHA, however, can be applied more or less

---

[46]Santoni de Sio and van den Hoven (n 1) 1.

[47]Jack Stilgoe and others, "Developing a Framework for Responsible Innovation" (2013) 42 Research Policy 1568, 1570.

[48]Santoni de Sio and van den Hoven (n 1) 12.

[49]Wagner (n 2) 115.

immediately onto existing ADM processes. For example, the time element could be fulfilled by ensuring that the organisational structure in which a task takes place allows that the human has enough time in order to conduct such task. The time element does not necessarily require that the ADM process is designed in a way that provides for sufficient time for the human. This means that the MHA concept has a short-term advantage over MHC. The latter cannot be applied directly onto existing systems, as it ties into a very early stage of the technology life-cycle, which means that currently deployed systems first need to be redesigned according to the tracking and tracing conditions.

One drawback of MHA is that each element is described in the form of a range, e.g., the more time the human has to conduct a task, the less likely a process is quasi-automated. That means that Wager does not set up a clear threshold as to when time is considered sufficient, which makes the assessment of whether a MHA element is fulfilled or not very vague and case dependent. Furthermore, most MHA elements contain terms that require further interpretation, since Wagner does not engage in a very detailed explanation of the seven MHA elements. One example is the the term *relevant information.*

3.4   Extended Meaningful Human Control

Due to the advantages and disadvantages identified in Section 3.3, this thesis suggests to combine the conditions of MHC and certain elements of MHA in order to define qualified human participation. This combined approach is referred to as Extended MHC concept. The analyses in Section 3.1.3 and Section 3.3 show that the tracking and tracing condition provide a sound structure to base a definition of qualified human participation on. The design approach of MHC poses a challenge to those engaging in its implementation. However, incorporating, and therefore technically cementing human control into an ADM system's architecture may hold great promise for effective human control over ADM processes, since the ADM process will be intertwined with, and dependent on human participation.

As Section 3.2.2 shows, certain individual elements of MHA constitute useful indicators for aspects that should be taken into account in order to maintain human autonomy in an ADM process. The advantage of MHA lies in the clear choice and indication of such aspects. Therefore, the time, qualification, support, and information element are adapted and used to complement and clarify the tracing condition. These MHA elements, when combined, can ensure that the moral and technical understanding required by the tracing condition, meets a certain quality standard.

The liability, adaption, and agency element are not incorporated into the Extended MHC approach. Due to the critique mentioned in Section 3.2.2, the suitability of the liability and agency element in order to explain qualified human participation is denied. The adaption element is not incorporated into the Extended MHC approach. That is because the tracking and tracing condition sufficiently provide for the aim of the adaption element. The way in which the two MHC are formulated requires the ADM process to be design around the human, i.e., that the system aligns with

relevant human reasoning and that its behaviour is traceable to human understanding. Furthermore, the adaption element does not provide any further clarification on how the adaption of the ADM process could be designed for.

The time element requires that the way in which the human achieves moral and technical understanding ensures that the human spends sufficient time to deliberate about technical and moral aspects of the system. Similar to the observations made in Section 3.2.2, enough time can prevent wrong understanding or abandonment of the training due to time pressure. This allows the human to whose understanding the behaviour of an ADM is traced back to to properly assess the system's capabilities and effects. If technical or moral training is conducted by merely sending an e-mail with a few lines, the time spent on such training is small. Sufficient time can also mean that training is provided frequently. Therefore, time frames for engaging in moral and technical understanding have to be defined. Other options are to evaluate the necessary time individually in an exchange with the human, which allows the possibility to adapt the time frame to personal traits of the human.

The qualification element is useful since it stipulates that a certain level of qualification is required. When combined with the tracing condition, the qualification element requires that the moral and technical training provided to the human has to lead to a sufficient degree of expertise and knowledge to understand the moral and technical implications of the ADM process. The qualification element could be ensured if, for example, the training is tied together with an examination.

Furthermore, the moral and technical understanding must be achieved in a way that allows the human to learn in a supported, effective and sustainable manner. If, as in the example in relation to the time element, training merely constitutes an e-mail with a few sentences, the human is not sufficiently supported to conduct and engage in the training. As discussed in Section 3.2.2, support mechanisms can take various forms, such as engaging with co-workers with similar tasks, or being able to discuss with supervisors. The type and the duration of such support has to be defined on a case-by-case basis, and in consultation with the relevant human.

The information element requires that the moral and technical understanding of the human has to flow from sufficient information. For example, consider a situation when two different approaches offer alternative solutions to technical problem. The human only knows one of these two approaches. To know both approaches, however, would be necessary in order to fully understand the technical problem. By only knowing one approach, the human is able to understand that there is a technical problem, and knows one way to solve it. However, by not knowing the alternative approach, the human lacks sufficient information to fully understand the issue at hand. In this example, having sufficient information enables the human to engage in critical thinking, hence actually understanding the problem and the solution to it. Furthermore, the information element means that the training has to be of scientific value, rather than be based on common knowledge. Therefore, the training offered to personnel should be held by experts in the relevant fields.

Therefore, the Extended MHC approach is formulated as follows:

1. Tracking condition: An ADM system must be programmed in a way that allows the system to be responsive and adaptable to relevant factual, legal and

moral considerations of a human agent in a given situation, i.e., the system should track the reasoning of the agent.[50] In order to provide for the tracking condition, the relevant considerations and the agents to be tracked by the system have to be defined. Furthermore, the degree of responsiveness of the system has to be established.[51]

2. Tracing condition: The actions of an ADM system must trace back to factual, legal and moral comprehension of at least one human that either designs or operates the system. This human must comprehend or be able to comprehend the capabilities and the possible effects of the use of the system. Further, the human must comprehend or be able to comprehend that the use of the system or its effects on the world lead to reactions from other human.[52] The moral and technical understanding of the human has to be characterised by sufficient time, qualification, support, and information.

3.5   Chapter Conclusion

Both the concept of MHC and MHA can be used as an interpretation of qualified human participation for the purposes of art. 22(1) GDPR in order to ensure that the human plays a decisive role in an otherwise automated process. However, the analyses of strengths and weaknesses of the two sets of requirements of MHC and MHA lead to the conclusion that the most valuable approach to defining qualified human participation is to combine the concepts to what is referred to as the Extended MHC concept. The design approach of MHC ensures that the possibility for human control over an ADM system is incorporated into the system's functioning. Furthermore, the tracking and tracing condition address and mitigate concerns of lessening of human responsibility and ensure human influence on the ADM process. The time, qualification, support and information element of MHA are used to complement and clarify the tracing condition.

---

[50]Santoni de Sio and van den Hoven (n 1) 6.
[51]Ibid 11.
[52]Ibid 8.

Chapter 4

Application of Extended Meaningful Human Control

In order to demonstrate how Extended MHC can aid to define qualified human participation for the purpose of art. 22(1) GDPR, this chapter applies the Extended MHC criteria to a GDPR-relevant case. Since the Extended MHC criteria have been defined in this thesis, it is not of value, and not the objective of the chapter, to assess whether the use case fulfils the Extended MHC criteria. The objective of this chapter is not to scrutinise, or to identify strengths and weaknesses of how the process of the use case is designed. Rather, the case serves illustrative purposes, and should show how Extended MHC performs in a real case.

The chapter will proceed as follows. First, the use case is introduced. Then, the Extended MHC criteria are applied to the use case. Lastly, the results are discussed.

4.1   Use Case: Parcoursup

The case that is investigated in this thesis is an application process called Parcoursup[1], which is a French online platform that serves the purpose of application and allocation for higher education study positions. Parcoursup is a centralised and automated procedure,[2] and constitutes the single official registration procedure that is available for higher education throughout France.[3]

The use of Parcoursup has been legally provided for by French law. Therefore, the exemption in art. 22(2)(b) GDPR might be applicable, which permits solely automated decisions if national legislation provides a legal basis for such. However, serious doubts have been raised in relation to the second condition of the exemption. That is, the relevant legislation must provide safeguards that ensure the protection of the data subject's rights, freedoms and legitimate interests. One example of such doubt is the lack of transparency. The educational institutions are not legally obliged to publish

---

[1]"Parcoursup" ⟨https://www.parcoursup.fr/index.php?desc=⟩ accessed 7 May 2019.

[2]Noëlle Lenoir and others, *Rapport au parlement du Comité Ethique et Scientifique de Parcoursup* ⟨https://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/194000051.pdf⟩ accessed 7 May 2019, 15.

[3]Ibid 15.

the selection criteria according to which applicants are selected. This fuels concerns regarding discrimination of prospective students based on factors such as place of residence and location of former high school.[4] Therefore, is questionable whether the Parcoursup application can benefit from the exemption of art. 22(2)(b) GDPR. Notwithstanding the outcome of the question whether art. 22 GDPR is applicable or not, the reason for choosing the Parcoursup use case does not only lie in its concrete GDPR-relevance, but in its suitability to showcase the Extended MHC approach presented in Section 3.4 to define qualified human participation. The Parcoursup assignment procedure goes through the following steps:

Phase 1: Main phase

Step 1: Applicants can access relevant information about the study programs (characteristics and local and national requirements for the program, required documents from applicants),prepare their application and submit a maximum of ten proposals for programs, without ranking them by order of preference.

Step 2: Each educational institution sets up a proposal examination commission ("Commission d'Examination des Voeux", CEV). The CEV has a list of applicants and access to the documents of the application files.The CEV assesses the submitted applications, and ranks them according to institution-internal selection criteria. The ranking resulting from the deliberations of the CEV is integrated on the platform. One or two CEV rapporteurs are appointed for each case, and the assessment is harmonised between the committee's examiners in order to avoid distortions. Each application is given an overall score based on the institutional-internal selection criteria.

Step 3: During the first admission round, Parcoursup sends admission responses to applicants. The order in which answers to proposals are sent to candidates is calculated for each institution by the Parcoursup algorithm.

Phase 2: Complementary phase

Step 4: After the first proposal round, applicants can enter the complementary phase, and express new wishes for programs that have not been filled yet. That is especially the case if the applicant has not yet received any proposal, or has not yet registered on the platform.

Step 5: The applicant can enrol in any of the proposed programs that are presented as "yes" or "yes, if" options ("yes if" means the possibility of enrolling for the desired program on condition that the applicant visits additional courses in order for the applicant to meet the required level).[5]

In order to assign students to programs to which the CEV has deemed the student eligible, an ADM process is deployed. That means that the assignment decision, i.e., whether or not the eligible student does in fact receive a proposal, is delegated to the Parcoursup ADM process. The ADM process takes three types of information into consideration: capacities of the institutions, the wishes of the applicants, and the rankings of applicants by the institutions. Based on these parameters, a match

---

[4]Jacques Toubon, *Décision du Défenseur des droits nº 2019-021* ⟨https : / / juridique . defenseurdesdroits.fr/doc_num.php?explnum_id=18303⟩ accessed 7 June 2019, 4.
[5]Lenoir and others (n 2) 21, 43-51.

is calculated.[6] This assignment task is conducted by an ADM system that processes personal data, such as name and address of the applicants. While the review of applications is conducted by humans, the distribution of study positions is subject to a decision of the ADM process of Parcoursup. Hence, art. 22(1) GDPR is applicable.

4.2 Parcoursup and Extended Meaningful Human Control

In the following sections, the Extended MHC concept is applied to Parcoursup. Each section lays out what the relevant criterion requires in relation to the use case.

4.2.1 Tracking Condition

The tracking condition demands that Parcoursup observes and implements relevant human factual, legal, and moral considerations in its allocation task. This alignment ability must be incorporated into the code of Parcoursup application.[7] The roadmap presented in Section 3.1.3 shows the necessary steps that have to be taken when the tracking condition is applied to a specific case.

Relevant Considerations to be Tracked by the ADM System

The tracking condition requires that the relevant considerations that should be tracked by the ADM system are identified.[8] One approach to identify such relevant considerations is to first define the task that the ADM process should fulfil. Once the task is identified, the reasoning to fulfil the task can be established.

In the Parcoursup case, the ADM process ranks prospective students, and allocates study positions based on this ranking. Section 4.1 points out that the Parcoursup ADM process takes into account three types of information in order to conduct its ranking task: Capacities of the institutions, the wishes of the applicants, and the eligibility assessment of the institutions.[9] These are all factors that are defined by either the educational institution or by the applicant.

Currently, the educational institutions are not legally required to publish their selection criteria.[10] That means that it is unclear what considerations the CEV follows during its eligibility assessment. In order to ensure that relevant considerations for the purposes of the tracking condition can be established and scrutinised, this lack of transparency has to be resolved. Therefore, the ministry of higher education should adopt relevant legislative and regulatory measures.

---

[6] Annika Joeres, *Parcoursup - Das Außerirdische Universitätsauswahlsystem der Französischen Regierung* ⟨https://algorithmenethik.de/2018/05/30/parcoursup-das-ausserirdische-universitaetsauswahlsystem-der-franzoesischen-regierung/⟩ accessed 20 May 2019.

[7] Filippo Santoni de Sio and Jeroen van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account" (2018) 5 Frontiers in Robotics and AI 1, 7.

[8] Ibid 7.

[9] Joeres (n 6).

[10] Toubon (n 4) 2.

Circumstances such as wealth, nepotism, race, gender, or other discriminatory considerations should not be relevant reasons to be tracked. Considerations flowing from these circumstances would render the ranking task flawed due to reliance on unfair or discriminatory motives. Instead, the appointed agent has to engage in, and Parcoursup has to track, considerations of fairness and non-discrimination, which then lead to unbiased results of the ranking task. For this purpose, relevant national legislation such as a anti-discrimination laws, guidelines published by the ministry of higher education, or similar documents should be tracked to ensure compliance with legal rules.

Appointment of Agency

Once the relevant considerations to be tracked are identified, the agent are suitable to engages in such reasoning have to be identified. Due to the relevant considerations pointed out in Section 4.2.1, possible agents are the personnel that designs Parcoursup, and the legislator or policy-maker that lay down the legal rules for the ADM process. For example, agency can be ascribed to a highly skilled designer of the system, who is able to critically assess and review the planning, error-correction and maintaining of the Parcoursup program. Such an agent is capable to engage in reasoning about the correct functioning of Parcoursup. Another suitable agent is a person involved in legal and ethical issues of Parcoursup, who knows and understands national legislation on issues such as anti-discrimination. This agent is capable to engage in reasoning on whether the outcome is morally and legally fair.

Amount of Responsiveness of the ADM System

Lastly, the Parcoursup application's ability to track relevant considerations has to be at a certain degree of proficiency. This means that a threshold regarding the quality and effectiveness of the responsiveness of the system has to be established.[11] Since the admission to a higher education institution is an important and potentially life-changing matter, the degree of responsiveness has to be set high. Therefore, the designers, the ministry of higher education, and other governmental institutions should establish a certain standard for the Parcoursup application.

4.2.2   Extended Tracing Condition

In order to fulfil the extended tracing condition, the actions of the Parcoursup application must trace back to factual, legal, and moral comprehension of a human that is programming or operating Parcoursup. The action of Parcoursup is the allocation of a study position. This assignment task must trace back to a moral and technical understanding of at least one human along the design and deployment chain of Parcoursup. In order to satisfy the tracing condition, the human must comprehend, or be able to comprehend, the capabilities and the possible effects of Parcoursup. Furthermore, the human must comprehend, or be able to comprehend,

---

[11]Santoni de Sio and van den Hoven (n 7) 7.

that the use of Parcoursup or its effects on the world lead to reactions from other humans.[12]

In order to understand the capabilities of Parcoursup, technical understanding of the ADM process is required. Due to their technical background, the staff that design and maintain the platform are likely to fulfil this condition. According to the analysis in Section 3.1.3, one possibility to enable non-technical staff to such technical understanding is that the educational institution or governmental agencies provide relevant technical training. Furthermore, the extended tracing condition requires awareness about Parcoursup's impact on society and the individual. Here, too, training should be offered both to technical and non-technical personnel in order to promote moral understanding and awareness of responsibility.

The time, information and support element further define what requirements such training has to fulfil.The time element requires that the training provided to staff provides for sufficient time to deliberate about technical and moral aspects of Parcoursup. This allows the human, to which the actions of Parcoursup are traced back to, to properly assess the system's capabilities and effects. The ministry of higher education should offer compulsory training regularly, and ensure that this training offers staff enough time to learn in a sustainable way.

The qualification element requires that the training provided guarantees that the human is sufficiently enabled to understand technical functionalities and moral implications of the system. Therefore, the training has to provide the human with relevant knowledge. In the Parcoursup case, this can be, for example, a sound understanding of the functioning of the Parcoursup algorithm. This understanding can be examined and ensured by conducting certification procedures or similar testing.

Furthermore, the training must be held in a form that allows the human to learn in a supported and sustainable manner. That means that the training offered to Parcoursup staff must enable such staff to engage with co-workers who have similar tasks. This can have the effect that work relationships are created that enable the staff exchange more thoroughly with each other in order to solve a problem in relation to the Parcoursup application. Furthermore, the training should be organised in an interactive way. This would enable discussions and more critical analysis with the matter of, for example, moral implications of the Parcoursup application. Supervisors should follow up on their staff after the training, and ensure, that the knowledge acquired during the training is put in practice. Another possibility to enable supported and sustainable understanding would be to offer additional specialised training to interested staff, such as lectures at universities, or attending conferences.

The information element requires that the human, whose technical and moral understanding the Parcoursup's action traces back to, has to have technical and moral understanding that emanates from sufficient information. In order to suffice the information element, the training offered to Parcoursup personnel could be held by external experts in the fields of computer science, sociology, ethics and law. The ministry of higher education could engage academics and practitioners to guide the

---

[12]Ibid 8.

training, or, as pointed out earlier, provide the opportunity to staff to take part in lectures at university.

4.3   Required Adaptation

There are no national laws that stipulate how to create the Parcoursup application.[13]. One way to implement Extended MHC is to adapt legislation, and to bind all parties involved in the Parcoursup application to provide for the tracking and extended tracing condition.

In order to fulfil the tracking condition, the design of Parcoursup has to be transformed. That means that a technical tracking ability has to be incorporated into the architecture of Parcoursup. The relevant consideration to be tracked, the humans agents, and the level of responsiveness have to be defined jointly by the designers, ethicists, educational institutions and governmental agencies. In order to ensure that relevant reasons for the assessment by the CEV are transparent, which would enable more effective scrutiny of the Parcoursup process, national legislation should also lay down a principle of publication of the institution-internal selection criteria.

The application of the extended tracing condition sets up high requirements regarding Parcoursup. To assess which human is suitable for the purposes of the tracing condition requires an in-depth analysis of a person's competences. In order to enable such dual-ability for technical and moral understanding, the French government, educational institutions and any other involved entity must provide relevant training to educate staff. Finally, the fulfilment of these requirements has to be monitored by the undertakings or state authorities that are involved the Parcoursup ADM process.

4.4   Chapter Conclusion

This chapter shows that the Extended MHC criteria can be applied to a GDPR setting. For the purposes of the tracking condition, the identification of relevant considerations to be tracked, the appointment of agency, and the level of responsiveness of the ADM process are steps that can be applied in the Parcoursup case. The extended tracking condition can also be applied to a GDPR setting, and prescribes that the involved entities ensure technical and moral knowledge of their staff. Both criteria demand a high level of engagement by entities deploying ADM processes, such as the provision of high quality training. Extended MHC is highly context-dependent, and therefore requires adaptation when applied to different settings. More testing of Extended MHC to other ADM processes should be done in order to further optimise its applicability.

---

[13]Lenoir and others (n 2) 117.

Chapter 5

Conclusion

Art. 22(1) GDPR states that a data subject has the right not to be subject to a decision that is based solely on automated processing, including profiling, which produces legal effects concerning him or her, or similarly significantly affects him or her. Chapter 2 illustrates that the wording, *a decision based solely on automated processing*, is imprecise and contradicts the title and rationale of the article. Therefore, this thesis suggests to interpret the formulation as *a solely automated decision.*

Another identified issue with art. 22(1) GDPR is that a narrow understanding of the term *solely* for the purpose of the article means that even a symbolic form of human participation in an ADM process is enough to comply with the article. In practice, this prevents the application of art. 22 GDPR. A broad understanding of the term should be adopted, which means that the degree of human participation that is required by art. 22(1) GDPR has to meet a certain threshold in order to evade the application of the article. This minimum threshold is referred to as qualified human participation.

In order to identify the requirements of qualified human participation, two human control concepts, MHC, as defined by Santoni de Sio and van den Hoven[1], and MHA, as defined by Wagner[2], are analysed in Chapter 3. Both concepts provide useful criteria to explain qualified human participation, but neither concept is able to fully define the term without being complemented by the other. Instead, this thesis suggests to combine the tracking and tracing condition of MHC with the time, qualification, support and information element of MHA. This combined approach is referred to as Extended MHC. The combined criteria demand that an ADM process is designed in a way that ensures that the ADM process aligns with relevant human reasoning, and traces back to relevant moral and technical understanding of a human. In order to meet the extended tracing condition, the human whose understanding the behaviour of an ADM process is traced back to, has to have sufficient time to engage in understanding, must be sufficiently qualified to understand, has received sufficient support when engaging in understanding, and has the opportunity to rely

---

[1]Filippo Santoni de Sio and Jeroen van den Hoven, "Meaningful Human Control over Autonomous Systems: A Philosophical Account" (2018) 5 Frontiers in Robotics and AI 1.

[2]Ben Wagner, "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems" (2019) 11 Policy & Internet 104.

on a sufficient amount of information to engage in understanding the ADM process.

The Extended MHC criteria clarify the EDPB and ICO's loose guidelines, and provide a set of clear requirements for the purpose of assessing qualified human participation, and therefore resolve legal uncertainty of the exact meaning of the term *solely* in art. 22(1) GDPR. Furthermore, this account of qualified human participation addresses the EU legislator's concerns that gave rise to art. 15 DPD, which is the predecessor of art. 22 GDPR. The Extended MHC criteria ensure that a human is not only present in the ADM process, but is able to actively influence the ADM process due to relevant reasoning and technical and moral competence. This means that the human that is actively participating in the ADM process is not placed there for symbolic reasons, but highly trained and aware of its responsibility for the behaviour of the ADM process.

The analysis in Chapter 3, and the successful application of Extended MHC in Chapter 4, show that it can be useful to apply concepts from other disciplines in order to interpret a legal article. The question of how interactions between human and machine should be shaped is an issue that does not pertain only to the GDPR. Therefore, a multidisciplinary, analogy-drawing approach to understanding qualified human participation is more helpful, and perhaps even necessary.

The Extended MHC concept requires research on how to program ADM processes that are able to engage in tracking of human relevant reasoning. That means that companies have to adapt current, and arrange future, ADM process design according to the Extended MHC criteria. Furthermore, companies that deploy ADM processes have to ensure that moral and technical training is provided, and successfully completed by staff that designs or operates the ADM processes. Ethicists should accompany technical staff in order to define, for example, the relevant reasons to be taken into account by the system, in order to suffice the tracking condition. Legal scholars should analyse questions such as liability in cases when Extended MHC is, or is not, established. Another question that requires further elaboration is whether the Extended MHC conditions should be cast in the shape of hard law or soft law, and how enforcement could be organised.

It is important to stress the limitations, and the consequences of these limitations, of the Extended MHC concept. First, the criteria require a case-by-case adaptation. The Extended MHC criteria point out what factors have to be adhered to, for example, a sufficient amount of time in order to engage in moral and technical training for the purpose of the extended tracing condition. However, how much time that in fact has to be granted is highly context-dependent, and therefore has to be assessed depending on the ADM process and human that is involved. Second, the Extended MHC criteria should be tested and evaluated in more use cases. Even though Extended MHC is intentionally formulated in a way that accommodates for a variety of settings, application to more use cases would scrutinise the concept's adaptability and effectiveness to establish qualified human participation for different settings.

The implementation of qualified human participation, and therefore the significance of Extended MHC, highly depends on whether and how art. 22 GDPR will be enforced. If scholars, national supervisory authorities, national courts, the EDPB,

and the CJEU refrain from resolving the ambiguities of art. 22 GDPR, the article will have no more impact than its predecessor. Therefore, the author encourages a more vivid discussion on the meaning of qualified human participation for the purpose of art. 22(1) GDPR.

Bibliography

Arkin RC, "Lethal Autonomous Systems and the Plight of the Non-Combatant" in R Kiggings (ed), *The Political Economy of Robots, Prospects for Prosperity and Peace in the Automated 21st Century* (Springer-Cham Palgrave Macmillan 2018) (citspace on p 20).

Article 29 Working Party, *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679* (WP251, 2017) (citspace on pp 1, 15, 16, 23, 27–29).

"Campaign to Stop Killer Robots" ⟨https://www.stopkillerrobots.org/⟩ accessed 18 May 2019 (citspace on p 20).

Centre for Information Policy Leadership, *Implementing and Interpreting the GDPR: Challenges and Opportunities, Towards a Successful and Consistent Implementation of the GDPR* (2016) ⟨https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_gdpr_amsterdam_workshop_report___2016.oct.06_.pdf⟩ accessed 2 April 2019 (citspace on p 11).

Chengeta T, "Defining the Emerging Notion of "Meaningful Human Control" in Weapon Systems" (2016) 49 NYU Journal of International Law 833 (citspace on p 19).

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 [1995] OJ L281/31.

European Commission, *Amended Proposal for a Council Directive on the protection of individuals with regard to the processing of personal data and on the free movement of such data* (COM (92) 422 final, 1992) (citspace on p 9).

European Commission, *Commission Communication on the Protection of Individuals in Relation to the Processing of Personal Data in the Community and Information Security* (COM (90) 314 final) (citspace on pp 8, 9, 24, 30).

European Data Protection Board, *Endorsement of GDPR WP29 Guidelines by the EDPB* (2018) ⟨https://edpb.europa.eu/news/news/2018/endorsement-gdpr-wp29-guidelines-edpb%5C_de⟩ accessed 2 May 2019 (citspace on p 1).

Finck M, "Smart Contracts as a Form of Solely Automated Processing Under the GDPR" (2019) 19-01 Max Planck Institute for Innovation & Competition Research Paper Series (citspace on p 12).

Joeres A, *Parcoursup - Das Außerirdische Universitätsauswahlsystem der Französischen Regierung* ⟨https://algorithmenethik.de/2018/05/30/parcoursup-das-ausserirdische-universitaetsauswahlsystem-der-franzoesischen-regierung/⟩ accessed 20 May 2019 (citspace on p 37).

Judgement of the German Federal Court VI ZR 156/13.

"Killer robots: UK Government Policy on Fully Autonomous Weapons" (*Article 36*, 2013) ⟨http://www.article36.org/wp-content/uploads/2013/04/Policy_Paper1.pdf⟩ accessed 20 May 2019 (citspace on p 19).

"Killing by Machine: Key Issues for Understanding Meaningful Human Control" (*Article 36*, 2015) ⟨http://www.article36.org/wp-content/uploads/2013/06/KILLING_BY_MACHINE_6.4.15.pdf⟩ accessed 20 May 2019 (citspace on pp 3, 20).

Lenoir N, Berry G, Dauchet M, Grenet J, Lucchesi L, and Moisan C, *Rapport au parlement du Comité Ethique et Scientifique de Parcoursup* ⟨https://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/194000051.pdf⟩ accessed 7 May 2019 (citspace on pp 35, 36, 40).

Loi n° 78-17 du 6 janvier 1978 relative à l'informatique, aux fichiers et aux libertés.

Matthias A, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata" (2004) 6 Ethics and Information Technology 175 (citspace on p 30).

Mendoza I and Bygrave LA, *The Right Not to Be Subject to Automated Decisions Based on Profiling* (Springer 2017) (citspace on pp 8–11, 24).

Paal BP and Pauly DA, *Datenschutz-Grundverordnung Bundesdatenschutzgesetz* (CH Beck 2017) (citspace on p 10).

"Parcoursup" ⟨https://www.parcoursup.fr/index.php?desc=⟩ accessed 7 May 2019 (citspace on p 35).

Pehrsson E, "The Meaning of the GDPR Article 22" (2018) 31 European Union Law Working Papers (citspace on pp 11, 12).

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1.

Report of the 2018 session of the Group of Governmental Experts on Emerging Technologies in the Area of Lethal Autonomous Weapons Systems ⟨https://undocs.org/en/CCW/GGE.1/2018/3⟩ accessed 29 January 2019 (citspace on p 20).

"Rights related to automated decision making including profiling" (*Information Commissioner's Office*) ⟨https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/rights-related-to-automated-decision-making-including-profiling/⟩ accessed 20 March 2019 (citspace on p 16).

Santoni de Sio F and van den Hoven J, "Meaningful Human Control over Autonomous Systems: A Philosophical Account" (2018) 5 Frontiers in Robotics and AI 1 (citspace on pp 3, 19–24, 29–31, 34, 37–39, 41).

Saxon D, "Autonomous drones and individual criminal responsibility" in E Di Nucci and F Santoni de Sio (eds), *Drones and Responsibility: Legal, Philosophical, and Socio-Technical Perspectives on the Use of Remotely Controlled Weapons* (Routledge 2016) (citspace on p 21).

Skitka LJ, Mosier K, and Burdick MD, "Accountability and Automation Bias" (2000) 52 International Journal of Human-Computer Studies 701 (citspace on p 26).

Stilgoe J, Owen R, and Macnaghten P, "Developing a Framework for Responsible Innovation" (2013) 42 Research Policy 1568 (citspace on p 31).

Toubon J, *Décision du Défenseur des droits nᵒ 2019-021* ⟨https : / / juridique . defenseurdesdroits.fr / doc_num.php ? explnum_id = 18303⟩ accessed 7 June 2019 (citspace on pp 36, 37).

United States Department of Defense, *Department of Defense Directive 3000.09: Autonomy in Weapon Systems, November 21, 2012* ⟨https://www.hsdl.org/ ?abstract&did=726163⟩ accessed 5 April 2019 (citspace on p 19).

Veale M and Edwards L, "Clarity, surprises, and further questions in the Article 29 Working Party draft guidance on automated decision-making and profiling" (2018) 34 Computer Law & Security Review 398 (citspace on pp 8, 9, 13).

Wachter S, Mittelstadt B, and Floridi L, "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation" (2017) 7 International Data Privacy Law 76 (citspace on p 14).

Wagner B, "Liable, but Not in Control? Ensuring Meaningful Human Agency in Automated Decision-Making Systems" (2019) 11 Policy & Internet 104 (citspace on pp 3, 19, 24, 25, 29–31, 41).

"What does the GDPR say about automated decision-making and profiling?" (*Information Commissioner's Office*) ⟨https://ico.org.uk/for-organisations/guide-to- data- protection / guide- to- the- general- data- protection- regulation- gdpr / automated- decision- making- and- profiling / what- does- the- gdpr- say- about-automated-decision-making-and-profiling/⟩ accessed 20 March 2019 (citspace on pp 15, 16).

Zarsky T, "Transparent Predictions" (2013) 2013 University of Illinois Law Review 1503 (citspace on p 10).