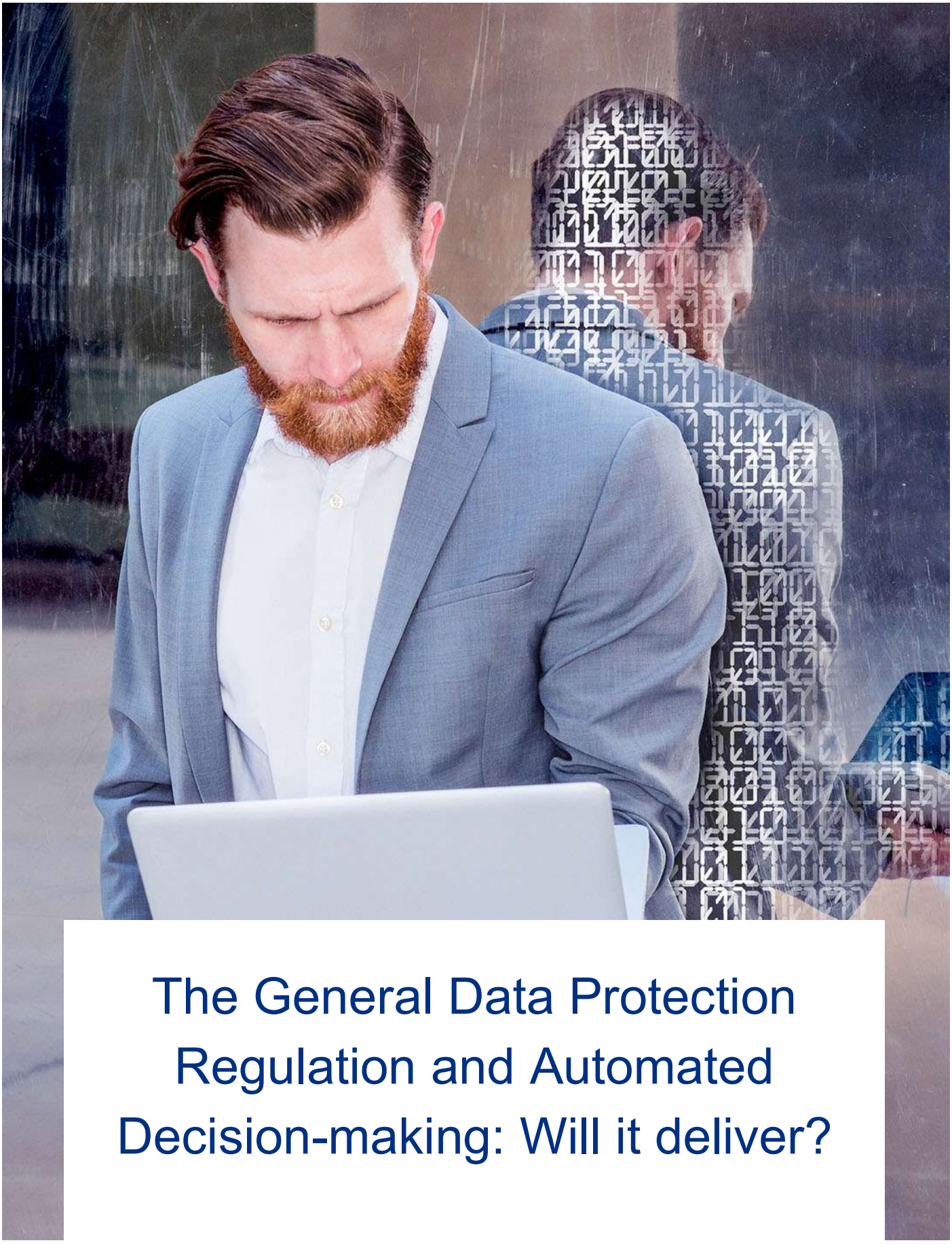
Discussion Paper Ethics of Algorithms #5



The General Data Protection Regulation and Automated Decision-making: Will it deliver?

BertelsmannStiftung

The General Data Protection Regulation and Automated Decision-making: Will it deliver?

Potentials and limitations in ensuring the rights and freedoms of individuals, groups and society as a whole

Stephan Dreyer and Wolfgang Schulz

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
www.bertelsmann-stiftung.org

Responsible for Content
Konrad Lischka
Ralph Müller-Eiselt

Authors
Stephan Dreyer, Wolfgang Schulz

DOI 10.11586/2018018 https://doi.org/10.11586/2018018

Contents

Preface 7

Summary 9

Executive Summary 11

1 Introduction 13

2 Potential risks of ADM systems and counter-measures aimed at safeguarding the interests of individuals, groups and society as a whole 14

3 GDPR requirements concerning ADM systems and relevant provisions of data protection law 17

3.1 Prohibition of pure ADM systems – with far-reaching exemptions 18

3.2 Information duties for the data controller of an ADM system *prior* to data processing 22

3.3 Conditions for user consent in the case of automated decisions 26

3.4 Rights of data subjects after data processing by ADM systems 27

3.5 Systemic and procedural duties of providers of ADM systems 30

4 Where the GDPR brings benefits: Approaches to safeguarding individual rights and freedoms 32

4.1 Transparency rules strengthen autonomy and personality rights of the individual user 33

4.2 Rights to obtain human intervention guarantee a ”human in the loop“ 33

4.3 Positive indirect effects of systemic and procedural duties 33

4.4 Role and measures of data protection authorities to safeguard the normative goals 34

4.5 Interim conclusions: GDPR’s starting points for measures safeguarding individual rights and freedoms 35

5 Where the GDPR falls short: Weak spots concerning group-related and societal interests 36

6 How the GDPR could bring benefits: Data protection approaches and instruments for the remaining risk potentials 38

6.1 Co-regulation: Certified codes of conduct as a support for commercial initiatives 38

6.2 GDPR: Expanding data protection authorities' regulatory options 39

6.3 Opening clauses: More restrictive national law requirements 40

7 Beyond the GDPR: Alternative regulatory tools not covered by data protection law 41

7.1 Explainability of automated decisions as a regulatory approach 41

7.2 Enhanced transparency and accountability provisions enabling third-party evaluations 42

7.3 Options for external evaluation without access to the system 43

7.4 Options for application of consumer protection and competition law for improved rectifiability 43

7.5 Options for application or adoption of regulatory tools from competition law and media law to ensure diversity 44

8 Conclusions 45

Literature 47

About the authors 49

Preface

Data protection cannot protect us from every imaginable danger. The fact that data are processed in conformity with data protection rules does not guarantee the quality of the conclusions that the software will deduce from them. When algorithmic systems evaluate data, risks are associated not only with the processing as such. It is above all the recommendations and decisions of software which constitute a threat to social inclusion and can, therefore, have an impact on individuals, social groups and society as a whole.

For example, in the United States and the United Kingdom, up to 70% of job applicants are evaluated and pre-selected by algorithmic decision-making systems before a human recruiter looks at the remaining candidates. In Germany, a number of companies are also beginning to make use of such automated processes and it seems likely that they will soon be used on a widespread basis. When compared with human decision-makers, these systems have an advantage in that they apply a decision-making pattern consistently to every case and are not prone to subjective distortion. Whereas human beings, for example, may allow themselves, at least subconsciously, to be influenced by the name or photograph of an applicant, software remains neutral.

However, algorithmic systems can also discriminate in a similar consistent manner. In 2014, the Wall Street Journal published an article about Kyle Behm, a student suffering from a bipolar disorder. He applied for temporary employment in retail, but, despite excellent testimonials and references, was not even invited to an interview by a number of employers. A friend who worked at one of these companies told him that he was a victim of the psychological tests conducted during the online application. The software had filtered out his application in all of the companies concerned. As a result of this process, Kyle and other applicants with a mental illness were unable to enter the labor market.

Unlike Kyle, the majority of applicants are left in the dark about why they have been rejected. And they seldom do anything about it. But if data subjects cannot understand automated decisions and if there are no complaints, it is impossible to reveal discrimination against individuals or groups.

For this reason, algorithmic systems that evaluate and assess human beings should be comprehensible for the individual and open to scrutiny with regard to systemic faults. Can data protection laws, and above all the new EU General Data Protection Regulation (GDPR) make a contribution in this regard? The application of the new regulation on May 25th, 2018 marks an important step toward harmonizing data protection throughout Europe and increasing data subjects' rights. The present report examines whether or not this will also be true with regard to automated decisions by software and algorithmic systems.

Wolfgang Schulz and Stephan Dreyer analyze in detail the relevant articles of the GDPR and elucidate whether the new regulation can promote more comprehensible and verifiable algorithmic decision-making systems. They examine to what extent the new regulation can safeguard individual interests such as personality rights on the one hand and societal goals such as non-discrimination and social inclusion on the other. In addition to this, they explore certain additions to the GDPR and alternative regulatory tools, which could complement its provisions. Such additional approaches are needed for algorithmic systems. The regulation has a restricted area of application and in view of its focus on safeguarding individual rights it is not capable of safeguarding group interests and societal values such as non-discrimination. In order to broaden the scope of what is possible within the framework of the GDPR, there is above all a need for a more active data protection supervisory authority which is prepared to look at the societal risks of ADM systems. Moreover, beyond the GDPR's scope, there is a need for additional approaches which can facilitate a more far-reaching evaluation and rectification of automated decisions. Besides provisions to open ADM systems to scrutiny by independent third parties, consumer protection law provides additional regulatory starting points.

We are publishing the present report as a working paper in order to make a contribution to a developing research area on which others can continue to build. We would be happy to receive ideas for enhancements and improvements, and, of course, constructive criticism. In order to facilitate a discourse of this kind, we are issuing the working paper under a Creative Commons license (CC BY-SA 3.0).

The analysis conducted by Wolfgang Schulz and Stephan Dreyer is part of the “Ethics of Algorithms” project in which the Bertelsmann Stiftung is taking a closer look at the societal consequences of algorithmic decision-making systems. A collection of international case studies (Lischka and Klingel 2017), an examination of the impact of algorithmic decision-making on social inclusion (Vieth and Wagner 2017), an analysis of the influence of algorithmic processes on societal discourse (Lischka and Stöcker 2017) and a paper on sources of error and responsibilities in algorithmic decision-making processes (Zweig 2018) have already appeared in the “Impulse Algorithmenethik” series.

Ralph Müller-Eiselt
Director
Program Megatrends
Bertelsmann Stiftung

Konrad Lischka

Summary

In algorithmic decision-making (ADM) systems, machines evaluate and assess human beings and, on this basis, make a decision or provide a forecast or a recommendation for action. This means that the data processing and the decisions it delivers contain *risks* for the users. On the one hand, there are individual rights such as informational self-determination being the core objective of data protection, personality rights and individual autonomy. On the other hand, there are group-related and societal interests such as fairness, non-discrimination, social inclusion and pluralism.

In order to attain these goals, experts have suggested the adoption of certain *measures* which contribute to making ADM processes transparent, individual decisions explainable and revisable, as well as to making the systems verifiable and rectifiable. Furthermore, ensuring the diversity of ADM systems can contribute to safeguarding the aforementioned interests.

Against this background, the present report focuses on the following *question*: To what extent can the EU General Data Protection Regulation (GDPR) and the new German Federal Data Protection Act (BDSG), both of which entered into force in May 2018, support such measures and protect the interests threatened by algorithmic systems? The analysis demonstrates that the Article 22 GDPR's scope of applicability with respect to ADM systems is quite restricted. In the few cases where the ADM specific provisions apply, it can to some extent create transparency and verifiability and thus help safeguard individual rights. However, regarding group-related and societal goals such as non-discrimination and social inclusion, the GDPR has little to offer. Discussing complementary regulatory tools beyond the GDPR is therefore necessary.

The present analysis gives a detailed account of why the scope of *application of the ADM specific provisions in the GDPR*, for a variety of reasons, is *closely restricted*. The GDPR prohibits only fully automated decision-making. Systems which prepare the basis for human decisions and give recommendations may still be used. For a prohibition to come into effect, an ADM system must make fully automated decisions on the basis of personal data, and the decisions made must have legal consequences or affect the data subject in a similarly significant way. If one of these three criteria is missing, the ADM-specific provisions of the GDPR do not apply. However, it is unclear in the case of ADM systems what a "decision" actually is and under what circumstances it produces "legal effects." Moreover, the regulation can hardly encompass the diversity of actual decision-making situations in which people consciously or unconsciously implement an automated decision or recommendation unquestioningly. Both the relatively narrow scope of application of the prohibition and the broad range of legal exceptions to the prohibition – first and foremost by consent given by the data subject – result in very limited cases in which an ADM system is actually prohibited. As a result, *(partly) automated decisions* are going to become *a normal part of our everyday lives*.

For ADM systems that are "exceptionally" permissible under the GDPR, the regulation contains legal provisions which can safeguard in part the *individual interests of the users*. Data controllers of ADM systems are subject to *transparency and information obligations* relating to the use of ADM systems in general as well as to the basic mechanisms of data processing and decision-making. However, the scope and depth of such transparency obligations are limited. It also remains unclear what the notion of "logic involved" actually means. Moreover, the regulation focuses on the data protection of the individual. For this reason, the scope and comprehensibility of the explanation of the "significance and the envisaged consequences" of the ADM decision is based on and limited by the perspective and cognitive skills of the average user. However, transparency provisions aiming at data subjects do not automatically lead to higher levels of basic rights protection in practical terms.

Regarding ADM systems, data subjects have a *right to disclosure* about the use of an ADM system in general as well as regarding the basic mechanisms of data processing and decision-making. Furthermore, they have the right to obtain human intervention. These rights, too, help safeguard individual rights and freedoms. They make it

possible to verify and – if needed – to overrule the automated decision. However, this does not constitute a right for data subjects or for independent third parties to scrutinize the entire system.

Systemic and procedural GDPR provisions regarding the design and implementation of ADM systems can help the data controller detect risks for the individual (and indirectly for groups) at an early stage and ensure minimum quality standards. These include privacy by design obligations, obligatory data protection impact assessments and binding corporate rules, as well as the appointment of a data protection officer. These regulatory tools have the potential to create a high level of awareness among data controllers regarding data protection issues which, in turn, can help safeguard individual rights and freedoms.

These controller-related duties can, in theory, be strengthened by the *data protection authorities*, who are granted encompassing disclosure and access rights. These authorities can scrutinize ADM processes and carry out impact assessments during data protection audits. However, once again, these audits focus only on the protection of individual rights.

Yet the GDPR does not offer great potential when it comes to protecting *group-related and societal interests* such as non-discrimination, social inclusion or pluralism. This would involve providing the option for an external inspection of the internal design of the ADM systems which would allow for the independent evaluation of its underlying concepts and processes. However, the GDPR transparency rules cannot facilitate such a deep insight into the system. Thus, it is not possible to uncover errors or misconceptions in the development and implementation of ADM systems as well as their potential effects on social interactions. Moreover, an overview of the actual diversity of ADM systems is difficult to acquire, given the context of system-related intransparency.

In order to protect group-related and societal interests as well as to improve system-related transparency and the external evaluation that this entails, there is a need for *complementary approaches*. Certain preventive measures *within the GDPR* can be strengthened for this purpose. For example, the data protection authorities could also ask for data protection impact assessments for all cases of ADM systems, including those that are not covered by Art. 22 of the GDPR. This would make it possible to identify risks at an early stage and to guarantee minimum protection standards. Furthermore, within the framework of the GDPR, the role of the data protection authorities could shift toward more public information and awareness building regarding potential societal problems, even if the authorities do not have enforcement powers that go beyond dealing with data protection violations.

However, beyond the scope of GDPR, *other regulatory tools* in practice have to be discussed if both supra-individual and societal goals are to be safeguarded. In order to improve the inspection of ADM systems, approaches that render a system more explainable can help improve the evaluation of ADM systems. Where such systems are already in use, enhanced transparency requirements could provide for better external assessment, for example, in the form of in-camera proceedings that protect the interests and confidentiality of the data controller. In order to rectify ADM systems already implemented, the use of regulatory tools found in competition and consumer protection law might be useful as they can result in faster forms of enforcement. The diversity of ADM systems can be supported by adopting regulatory tools provided by cartel law. Furthermore, media law requirements could foster pluralism among ADM systems that affect information access and have some impact in influencing public opinion. This would allow alternative regulatory approaches to protect supra-individual interests not covered by the GDPR, which focuses primarily on safeguarding individual rights and freedoms.

Executive Summary

In algorithmic decision-making systems (ADM systems) machines evaluate and assess human beings and, on this basis, make a decision or provide a forecast or a recommendation for action. Thus not only the data processing as such, but above all the decision that results from the processing contains *risks* for the user. On the one hand there are individual rights such as informational self-determination as the scope of protection directly covered by data protection, personality rights and individual autonomy. On the other hand there are group-related and societal interests such as fairness, non-discrimination, social participation and pluralism.

In order to attain these goals, experts have suggested the adoption of certain *measures* which contribute to making ADM procedures transparent, individual decisions explainable and revisable, as well as to making the systems verifiable and rectifiable. Furthermore, ensuring the diversity of ADM systems can make a contribution to safeguarding the mentioned interests.

Against this background the present report focuses on the following *question*: To what extent can the EU General Data Protection Regulation (GDPR), which applies from May 2018, and the new German Federal Data Protection Act (BDSG), which entered into force at the same time, support such measures and protect the interests threatened by algorithmic systems. The analysis demonstrates that the GDPR's room for manoeuvre in the area of ADM systems is quite restricted. In the few cases the regulation applies to it can to some extent create transparency and verifiability and thus help to safeguard individual rights. However, regarding group-related and societal goals such as non-discrimination and participation the GDPR has little to offer. For this reason there is a need to discuss complementary regulatory tools beyond the GDPR.

The present analysis gives a detailed account of why the *application of the GDPR* to ADM systems, for a variety of reasons, is *closely restricted*. The GDPR prohibits only fully automated decision-making. Systems which prepare the basis for human decisions and give recommendations may still be used. For the prohibition to come into effect, ADM systems must make fully automated decisions on the basis of personal data, plus the decisions must have legal consequences or similarly affecting the data subject significantly. If one of these three criteria is missing, the ADM-specific provisions of the GDPR do not apply. However, it is unclear in the case of ADM systems what a "decision" actually is, and under what circumstances it produces "legal effects". Moreover, the regulation can hardly encompass the diversity of actual decision-making situations in which people consciously or unconsciously implement an automated decision or recommendation unquestioningly. Both the relatively narrow scope of application of the prohibition and the broad range of legal exceptions to the prohibition - first and foremost by consent given by the data subject – result in very limited cases in which an ADM system is actually prohibited. Thus *(partly) automated decisions* are going to become a *normal part of our everyday lives*.

For ADM systems that are "exceptionally" permissible under the GDPR the regulation contains legal provisions which can partly safeguard the *individual interests of the users*. Data controllers of ADM systems are subject to *transparency and information obligations* relating to the use of ADM systems in general as well as to the basic mechanisms of data processing and decision-making. However, the scope and depth of such transparency obligations are limited. It also remains unclear what the notion of "logic involved" actually means. Moreover, the regulation focuses on the data protection of the individual. For this reason the scope and comprehensibility of the explanation of the "significance and the envisaged consequences" of the ADM decision is based on and limited by the perspective and cognitive skills of the average user. However, transparency provisions aiming at data subjects do not automatically lead to higher levels of basic rights protection in practical terms.

Regarding ADM systems data subjects have a *right to disclosure* about the use of an ADM system in general as well as regarding the basic mechanisms of data processing and decision-making. Furthermore, they have the right to obtain human intervention. These rights, too, help to safeguard individual rights and freedoms. They make it possible to verify and – if needed – to overrule the automated decision. However, this does not constitute a right for data subjects or for independent third parties to scrutinize the whole system.

Systemic and procedural GDPR provisions regarding the design and implementation of ADM systems can help the data controller to detect risks for the individual (and indirectly for groups) at an early stage and to ensure minimum quality standards. These include privacy by design obligations, obligatory data protection impact assessments and binding corporate rules, as well as the appointment of a data protection officer. These regulatory tools have the potential to create a high level of awareness with the data controller regarding data protection issues, and thus helping to safeguard individual rights and freedoms.

These controller-related duties can in theory be strengthened by the *data protection authorities*, who are granted encompassing disclosure and access rights. They can scrutinize ADM processes and carry out impact assessments during data protection audits. However, the focus of these audits is only the protection of individual rights, once again.

Yet the GDPR does not offer great potential when it comes to protecting *group-related and societal interests* such as non-discrimination, participation or pluralism. A prerequisite for this would be the option for an external inspection of the internal design of the ADM systems in order to be able to evaluate independently its basic concepts and processes. However, the GDPR transparency rules cannot facilitate such a deep insight into the system. Thus it is not possible to uncover errors or misconceptions in the development and implementation of ADM systems as well as their potential effects on social interactions. Moreover, an overview over the actual diversity of ADM systems is difficult to acquire against the background of system-related intransparency.

In order to protect group-related and societal interests as well as to improve system-related transparency and external evaluation, there is a need for *complementary approaches*. For this purpose certain measures *within the GDPR* can be strengthened. For example, the data protection authorities could also ask for data protection impact assessments for all cases of ADM systems, including those that are not covered by Art. 22 GDPR. This would make it possible to identify risks at an early stage and to guarantee minimum protection standards. Furthermore, within the framework of the GDPR the role of the data protection authorities could shift towards more public information and awareness building regarding potential societal problems, even if the authorities do not have enforcement powers that go beyond dealing with data protection infringements.

However, beyond the scope of GDPR *other regulatory tools* in practice have to be discussed in order to be able to safeguard both supraindividual and societal goals. In order to improve the inspection of ADM systems, certain approaches can contribute to the greater explainability of the systems. In case such systems are already in use, enhanced transparency requirements could provide for better external assessment, e.g. in the form of in-camera proceedings that protect the interests and confidentiality of the data controller. In order to rectify ADM systems already implemented, it seems possible to use regulatory tools from competition law and consumer protection law, since they might result in faster forms of enforcement. The diversity of ADM systems can be supported by adopting regulatory tools provided by cartel law. Furthermore, media law requirements could contribute to pluralism in the case of ADM systems that have an effect on information access and that influence public opinion formation. This way, alternative regulatory approaches can protect supraindividual interests that the GDPR does not cover, since it focuses primarily on safeguarding individual rights and freedoms.

1 Introduction[1]

Interest in software systems that evaluate people and make decisions based on algorithms – algorithm decision-making (ADM) systems – is growing primarily as a result of the potential impact such decisions can have on individuals, groups and society as a whole. In several publications and at various events and meetings, the Bertelsmann Stiftung has identified potential problems and risks associated with ADM systems affecting individual and societal goals while developing a series of practical measures designed to counteract these effects (Bertelsmann Stiftung 2017a). When the EU General Data Protection Regulation (GDPR)[2] and the completely overhauled version of Germany's Federal Data Protection Law (BDSG n.F.)[3] both became applicable on 25 May 2018, new data protection regulations have applied that contain specific provisions relevant for the processing of personal data in automated decision-making systems. This report thus identifies the extent to which these new regulations can constitute a basis for measures designed to minimize the risks associated with ADM systems and to protect the interests of individuals, groups and society as a whole. In addition, this report explores whether additional regulations in data protection law or in other areas of the law might be necessary and helpful.

This report is divided into four sections. Chapter 2 briefly outlines the risks posed by ADM systems to individuals, groups and society as well as measures designed to safeguard their interests. Chapter 3 describes the GDPR and BDSG n.F. regulations targeting ADM systems as well as the measures designed to protect individual, group and societal interests. In view of the need for action, Chapter 4 examines the legal framework in order to determine the extent to which data protection norms can provide relevant provisions in mitigating risks for individuals, groups and society as a whole. Chapter 5 identifies particularly thorny issues associated with ADM systems that neither the GDPR nor BDSG can currently resolve. The report concludes with a survey of alternative management approaches and tools within (Chapter 6) and beyond (Chapter 7) the data protection framework which might prove effective in addressing currently unmet needs.

[1] We would like to thank Florian Wittner, Florian Riebe, Sonja Lübeck, Lumnie Izeti and Johanna Stelling for their contributions to this report.
[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), OJ L 119, 4.5.2016, p. 1–88.
[3] BDSG i.d.F. des Datenschutz-Anpassungs- und -Umsetzungsgesetz EU – DSAnpUG-EU, 30. 6. 2017, BGBl I pp. 2097.

2 Potential risks of ADM systems and counter-measures aimed at safeguarding the interests of individuals, groups and society as a whole

Automated decisions can deliver benefits for individuals and society as a whole. The use of imaging processes in automated medical diagnosis, for example, clearly illustrates the positive potential inherent to this kind of technology. Similarly, in the field of environmental protection, the use of automated decision-making is effective in managing water quality and reservoirs or forecasting floods and air pollution. However, when it comes to regulatory frameworks for automated decision-making, legal practitioners and policymakers are primarily concerned with the risks ADM poses to legal rights and interests. The regulations they develop are thus designed to counteract these risks. The present synopsis of the legal, scholarly and societal debates, ideas and proposals regarding ADM systems thus focuses primarily on the threats to and risks for individual rights and freedoms as well as group and societal goals.

Risks can emerge in each phase of developing an ADM system and embedding it into a societal context (see also Zweig 2018). In the design phase, for example, there is the risk of selecting the wrong analytical concept or mathematical model, or that irrelevant data (for the task at hand) is (pre)selected.

- Errors in operationalizing social phenomenon and ideas – or applying aspects thereof irrelevant to resolving the targeted problem – can produce risks. Measures designed to render the selection of concepts or applied criteria transparent and identify relevant modes of operationalization can help resolve these issues.
- The ex-ante assessment of the social impact of ADM systems can be based on mispredictions. Subsequent evaluations that audit content can result in incorrect assessments.
- Programming that does not account for potential use in other operational areas that may be relevant for basic rights can render ex ante assessments obsolete. Impact assessments should therefore be conducted whenever an algorithm is used in a different area of application.
- The risk that the design logic will prioritize short-term economic efficiency gains over societal suitability can be mitigated if common good interests are considered in development processes (e.g., identifying a suitable logic model).

As the process continues, the system is embedded in a social context (implementation phase). This leads to certain risks as the developed system comes into practical use. The training data selected at an earlier stage and the system that has been trained on this basis are now subject to real-life conditions.

- When an ADM system is up and running there can be a self-reinforcing kind of feedback in which an element of distortion (or “bias”) that is already contained in the selected data is exacerbated. Such algorithmic bias cascades derive from the interaction of data selection and the systems that learn on this basis. It has been suggested that these risks might be addressed by demonstrating the falsifiability of the system decision. This can be done by making human verification and rectification of a decision possible and through a critical evaluation of the database and data structure (Barocas/Selbst 2016).
- There is a risk that statistics-driven decision-making systems will not sufficiently depict special situations and individual issues (“outliers”), which can result in structural disadvantages for data subjects. Individual single-case assessments that include the possibility of subsequent rectification can address this risk.
- Depending on their design, implemented systems can be black box systems involving non-transparent decisions. Non-human assessments of human beings and the helplessness associated with the “fear of being categorized” by a machine can be dealt with by insisting on the verifiability and scrutiny of a system’s logic and nature, thereby resulting in a better understanding of the opportunities and risks associated with an ADM system (Keats Citron/Pasquale 2014).

Moreover, explainable automated decisions can provide better insight when it comes to taking individual, group and societal interests into consideration.

- Risks can also materialize whenever the results of ADM systems are used for other purposes, for example, when aggregated data analyses are given to third parties. This can be prevented by stating that the results are to be used for a specific purpose only.
- In the area of the automated aggregation and organization of media content by intermediaries, network effects can lead to a consolidation of market power and a distorted prioritization of content, thereby fostering new means of influencing public opinion. Transparency instruments and diversity provisions are the primary tools under discussion for various types of prioritization logic (Schulz/Dankert 2016).

Other identified risk areas derive from the socio-technical consequences of the use of ADM systems (impact phase).

- The danger of undesirable social effects resulting from the use of ADM systems can be reduced by conducting comprehensive evaluations, particularly those that examine ADM impact on specific social contexts. Any review of a system should make a point of examining unforeseen social developments and the socioeconomic aspects of advantages and disadvantages, and risks and benefits.
- If algorithm monocultures begin to materialize (for example, when specific ADM architectures and logics gain predominance in a certain area), other and possibly better approaches may be suppressed. Underdeveloped or lack of innovation would lead to the widespread application of a handful of ADM processes and procedures. Data subjects affected “only” by the decisions of an individual system would in this case be held back systematically from making full use of their opportunity to participate. This negative result could be prevented by adopting measures designed to ensure the diversity of ADM processes.

The identified problem areas harbor risks in terms of creating and perpetuating social inequality, the unequal treatment of equals, and the dehumanization of decision-making, which involves rendering individuals the object of mathematical calculations. It is worth noting that this involves individual-based issues such as freedom of action, personality rights and fairness as well as group-relevant issues such as non-discrimination. In addition, these risk areas are relevant for society as a whole, particularly in terms of diversity. Another normative goal relevant for each level (i.e., individual, group, societal) is social inclusion. Risks to inclusion that may result from ADM systems can manifest themselves in a variety of ways both for individuals, for groups (e.g., minorities) and society as a whole. It is therefore not possible to clearly assign the goal of non-discrimination to only the individual level, the group-related level or the societal level. Some of the goals – in particular autonomy, personality rights and non-discrimination – are already addressed by legal norms and constitute basic elements of fundamental and human rights in national, European and even international frameworks. Other goals such as fairness, social inclusion and diversity, though they express broadly shared values and are the subject of ethical debates, do not presently constitute legal norms. This distinction is important to bear in mind when considering whether the state has a (legal) duty to take action, particularly when this difference is not always kept up in public discourse.

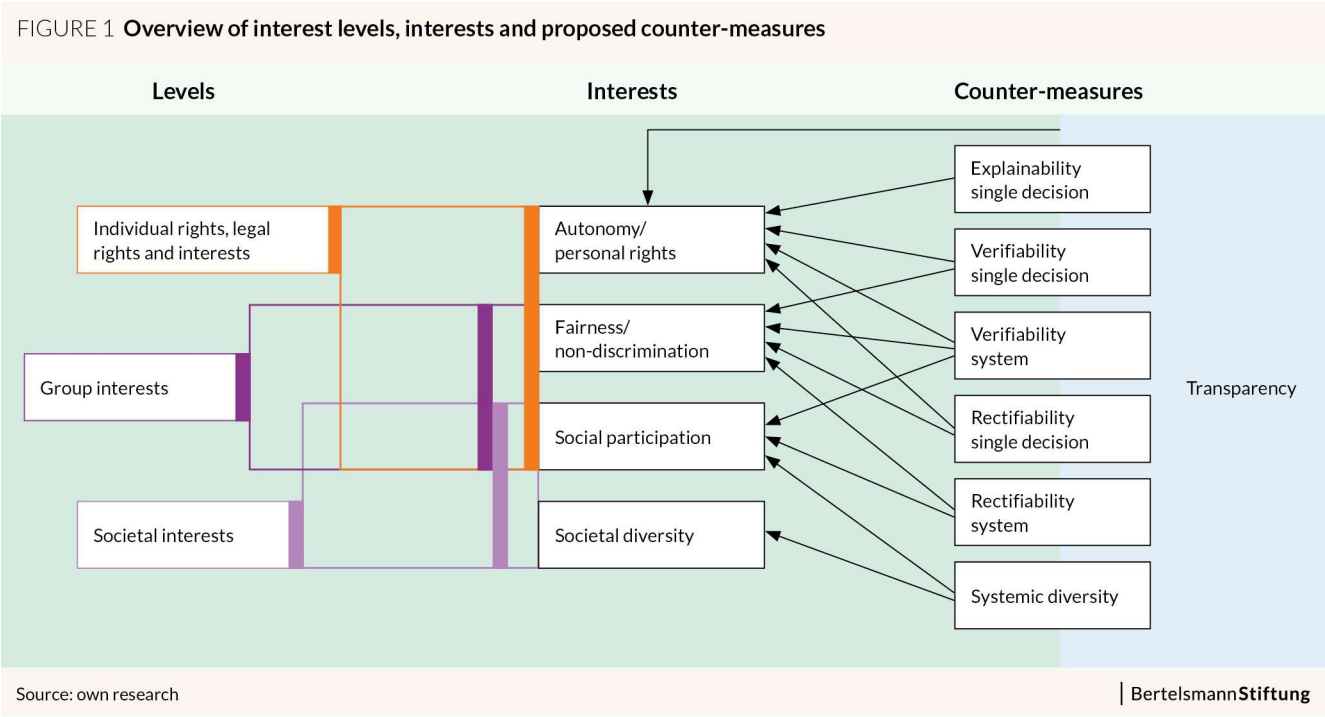
The individual tools recommended here as a response can be combined in policies designed to ensure transparency, the explainability of individual decisions, the verifiability of an individual decision and the process, the rectifiability of an individual decision and process and, finally, to guarantee the diversity of ADM systems.

Measures designed to ensure transparency are often deemed to be a panacea for all ills. Publishing information on ADM processes is supposed to make it easier for individuals to understand what they are and do while also, ideally, allowing the public to identify and criticize erroneous assumptions, ideas and decisions. In addition to transparency, measures related to single-case decision-making make a significant contribution to safeguarding autonomy and personality rights. This involves regulations that ensure the explainability, verifiability and rectifiability of a specific automated decision. Individual interests such as fairness and freedom from discrimination can also be safeguarded with the help of verifiable and rectifiable individual decisions. However, since discriminatory ADM

systems regularly effect specific ethnic groups, it seems necessary that we protect non-discrimination by ensuring systemic verifiability and rectifiability of the whole automated decision-making process. Explainability, verifiability and rectifiability all presuppose some knowledge of the ADM system. This makes transparency a counter-measure with the potential to have a positive effect on all of the aforementioned goals. Transparency can have a direct positive impact on the autonomous exercise of the right to informational self-determination, which is a manifestation of personality rights. And it can have an indirect positive impact as a vehicle for the improved explainability and verifiability of individual decisions and of the process as a whole.

In order to ensure social inclusion for individuals and society more generally, protecting systemic diversity, in addition to the verifiability and rectifiability of procedures that have already been implemented, is important for reducing systematic (and virtually unavoidable) exclusion risks that result from the practical one-sided use of certain ADM processes. Moreover, safeguarding ADM diversity is of central importance when it comes to maintaining societal diversity in automated decision-making systems.

Fig. 1: Overview of interest levels, interests and proposed counter-measures



The overview of identified goals and the measures proposed to safeguard them form a complex network (see Fig. 1). In what follows, this overview serves as the basis of our legal analysis. The report seeks to clarify whether and to what extent the data protection regulations in GDPR and BDSG n.F. provide the instruments necessary to aptly protect the identified individual, group and societal interests.

3 GDPR requirements concerning ADM systems and relevant provisions of data protection law

An examination of data protection law is especially pertinent when dealing with tools designed to safeguard the interests affected by ADM systems, as it already contains explicit rules for algorithmic systems. The EU General Data Protection Regulation (GDPR) and the new Federal Data Protection Law (BDSG n.F.), both of which became applicable in May 2018, contain, in addition to general regulations pertaining to data-processing providers, specific regulations relating to systems that make automated decisions on the basis of processing personal data. The following chapter identifies those provisions of the data protection law framework that directly or indirectly envisage measures designed to safeguard the goals alluded to above. It describes their area of application and scope as well as possible obstacles and difficulties associated with their implementation.

The fact that rather different goals have to be safeguarded does not immediately rule out an examination of data protection law. The ADM-specific requirements mention a number of different protected goods. The purpose of the GDPR is defined in Art. 1 (2), which states that it "protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data." Thus, the central idea is to safeguard the fundamental right specified in Art. 8 of the EU Charter of Fundamental Rights, which covers the protection of personal data. The secondary goal – to safeguard all basic rights and freedoms – points to the role of data protection as what is sometimes called run-up law; it protects central immaterial positions such as the autonomy of action and personal rights (in Germany these include informational self-determination), not only on account of these rights, but in order to counter a second wave of threats to the exercise of other rights and freedoms that emanate from data processing (Bull 2006).

Data processing poses a risk to a number of basic rights. In particular, there can be various kinds of individual and group discrimination, which can result in infringements on the exercise of other basic rights such as freedoms of conscience, religion and expression. Nevertheless, the focus of GDPR lies on data protection as stipulated in Art. 8 EU Charter of Fundamental Rights. It is thus not always possible to assume that other important rights and values – such as those alluded to above – will automatically be protected by the regulation. In certain cases, these rights can in fact protect interests (e.g., freedom of information) that even run counter to data protection. In taking its bearings from the decision-making consequences of automated decisions, EU legislators have departed from the straightforward protective purpose of data protection and have aimed above all to protect against the deleterious effects of decisions in the wake of automated data processing. This shift of the focus of protection from data processing to decision-making consequences is remarkable, because in this way the classical regulatory tools of data protection are being harnessed for the attainment of rather different protective purposes, even though an assessment of the efficacy of these regulatory tools has not been undertaken. It is also remarkable that neither within the framework of the legislative procedure nor in the legal commentaries there has hitherto been any systematic examination of this change of paradigm. To some extent, the two terms – data processing and automated decision-making – are used as if interchangeable. The significance of the distinction between decision-relevant data processing on the one hand and the decision itself and its consequences on the other must still be discussed in the context of the protective purpose of data protection (see below, chapter 3.1.1 and chapter 3.2.1).

The fact that the GDPR and the new BDSG provisions are public but have not been applied at the time of conducting this analysis should be kept in mind while considering what follows below. The statements and assessments made here on the basis of the published legal commentaries should play an important role in forthcoming administrative decisions and court rulings. Until then, legal difficulties in the interpretation of the (sometimes vague) legal concepts will mean legal uncertainty for data controllers and data subjects alike.

A consideration of the safeguarding of individual, group and societal interests by individual data protection tools should first address the question of the basic legal admissibility of ADM systems as specified in the GDPR.

3.1 Prohibition of pure ADM systems – with far-reaching exemptions

If automated decisions were generally inadmissible, the risks of ADM systems could not in fact materialize in the area covered by the GDPR. In fact, the GDPR lays down the principle that a person has "the right not to be subject to a decision based solely on automated processing" (Art. 22 (1) GDPR). The wording of the regulation defines this principle as the right of the individual to permit the data controller to make use of systems of this kind. However, in light of the subsequent structure of the regulation – paragraph 1 describes the principle, paragraph 2 the exceptions, paragraph 3 the special requirements for systems not covered by paragraph 1 – it seems to be a prohibition. Basically the use of such ADM systems is not permissible. This principle is not new. Article 15 of the EU Data Protection Directive and the present § 6a BDSG contain similar provisions. The two norms have hitherto played only a minor role in legal practice (but cf. Bygrave 2001). The political discussions in the context of GDPR deliberations have emphasized the significance of these norms in the face of increasing automated decision-making on the basis of personal data (Zarsky 2017).

However, the scope of the prohibition as specified in Art. 22 (1) GDPR is clearly restricted. The automated decision must first be made on the basis of the processing of personal data. Cases where the data is not related to a person – perhaps it has been anonymized or it was never associated with human being – do not fall within the scope of application of the regulation in general is, cf. Art. 2 (1) GDPR. Also if the data being processed are related to individuals other than the system user, it is not an ADM system as specified in the regulation.

Moreover, Art. 22 (1) of the GDPR refers only to ADM systems in which the decision is made "solely on automated processing." Only systems that make decisions without any human intervention are prohibited. The GDPR does not state that any kind of automated decision or data processing on which it is based is not permissible. Thus, the system must first make a decision. The prohibition does not apply to systems that "only" process personal data in automated form for purposes of preselection and preparation for decision-making, or to recommendation systems in which a human being makes the final decision ("decision support systems" or DS systems). The prerequisite is that the participation of a human being within the decision-making architecture is not merely a formal act. The decision-making individual must in fact be in a position to make his or her own decision on the basis of the actual data, even if it runs counter to the automated recommendation. The latter plays a central role with regard to the right to obtain human intervention (see chapter 3.4.1). To what extent the person indeed has decision-making leeway can be monitored by the local data protection authority within the framework of an inspection (on such supervisory powers, see chapter 4.4). At this juncture, what turns a single processing step into a "decision" in the legal sense of the word remains an open question in view of the numerous ways of constructing different kinds of decision-making architectures in which an ADM system and human action each play a part.

Lastly, the decisions made by an ADM system must produce "legal effects" concerning the data subject or has to similarly affect him or her. Legal effects can be assumed whenever a decision leads to a change in the legal status of the person concerned, for example, the denial of a legal entitlement or the issuance of a negative administrative order. To what extent the provisions can also be applied to cases in which the decision of the ADM system merely generates a legal advantage for the person concerned is a contentious issue. In view of the data processing also needed in this case, the protective purpose of the GDPR seems to come into play, especially since at the time when the data is processed, the result of the decision-making is not known. In the framework of private contracts to which the principle of private autonomy applies, a negative decision with regard to the conclusion of a contract or a contractual commitment under special circumstances does not qualify as legal impact. In such cases, the wish of the data subject to conclude a contract is not granted, yet his legal position is not altered. However, a "considerable adverse effect" may well be involved (see below). The situation varies with regard to the treatment of decisions that lead to the conclusion of a contract. Here the data subject enters into a new legal position with regard to the data controller and Art. 22 (1) GDPR applies.

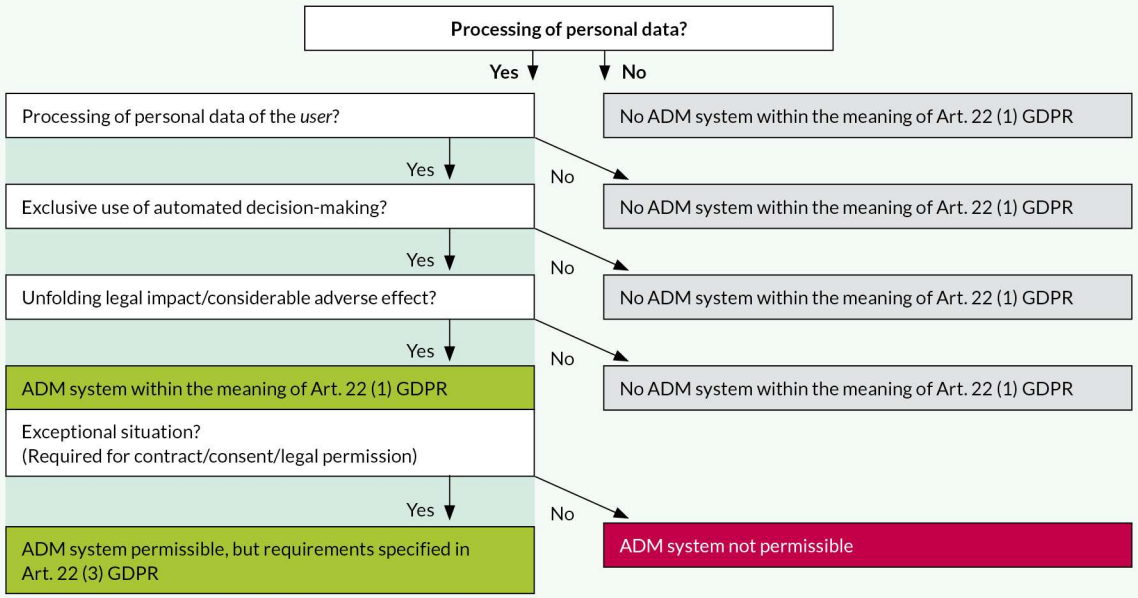
A specific answer to the question of when a data subject is a “similarly affected” can – in contradistinction to the legal effect – be seen in cases where a data subject outside his or her legal entitlement is completely denied the conclusion of a contract, or offered one on worse terms. A negative impact can also be assumed to exist whenever the decision is not exclusively disadvantageous in legal terms, for example, when a contract is concluded with the data subject but at a higher price or with incriminating conditions attached. Effects that do not have legal consequences, for example, decisions about the online advertising content that can be shown to a user on the basis of profiling measures, can be comprised by this. At what stage one reaches the level of “similarly affected” has not yet been specifically discussed. A central feature in the evaluation of what “similarly” actually means is the economic or practical significance of the decision-making object and the durability of the adverse effect. On the other hand, consequences of an automated decision that are merely troublesome or felt to be inconvenient should not be deemed to constitute a similar effect.

This isolated-case perspective cannot take into account the degree of damage that comes from the repetitive use of automated decision-making systems. If numerous systems deployed for the same purpose (e.g., profiling) repeatedly make decisions about an individual in the same way, this might result in a systematic impairment over a long period of time and with a large number of decisions. With regard to the broadly formulated protective purpose of the GDPR, which also encompasses the prohibition of discrimination specified in Art. 21 of the EU Charter of Fundamental Rights, such (discriminatory) decisions can theoretically become the subject of a GDPR examination. As long as there is no terminological precision with regard to what “similarly affected” actually means in practical legal terms, the legal debate will for the time being be confined to the framework of examinations conducted on a single-case basis.

The principle of the prohibition of pure ADM systems is being softened in practice, not only through the restrictive scope of application of the prohibition as specified in Art. 22 (1), but also by the far-reaching exceptions specified in Art. 22 (2) GDPR (cf. Mendoza/Bygrave 2017). Automated decision-making systems are thus “exceptionally” permissible if: (a) the decision is necessary for entering into, or for the performance of, a contract between the data subject and a data controller; (b) the ADM decision is authorized by the laws of the controller’s country; or (c) the decision is based on the data subject's explicit consent. Lastly, in cases of the exceptional permissibility of automated decision-making systems, Art. 22 (3) GDPR makes specific demands (see chapter 3.1.4).

Fig. 2: Decision tree on the scope of application of Art. 22 (1) of the GDPR

FIGURE 2 **Decision tree on the scope of application of Art. 22 (1) of the GDPR**



Source: own research | BertelsmannStiftung

3.1.1 Exemption 1. ADM is necessary for entering into or performance of a contract

The use of pure ADM systems is permissible in cases where the automated decision is required for the conclusion or the performance of a contract. However, the automated decision does not have to be the subject of the contract, and can, in fact, be no more than a decision-making basis for a conclusion of a contract. Here the GDPR provides a legal exemption to the prohibition of simple ADM systems especially where the recitals of the regulation see a major application of ADM systems. For example, Recital 71 mentions online credit applications, which are referred to in Art. 22 (1) GDPR. But these are then permitted via the exemption in Art. 22 (2) a) GDPR.

The problematical aspect of this exemption is that in cases where there is an imbalance between data controllers and data subjects, the data controller has a structural advantage when it comes to formulating the purpose and the various parts of the contract. Through such a one-sided definition of the purpose of the contract the data controller can also define the reasons for a (purported) necessity and can include the necessity of automatic decision-making in the contract on his own. Yet the actual necessity can become the object of legal inquiry. The origins and the reasons for the necessity has to be comprehensible from the perspective of neutral observers and has to be specifically focused on forms of automated decision-making systems, e.g. because entering the contract is subject to numerous items based on mathematical calculations and the resulting automated decision is directly connected with the conclusion of the contract. An alternative viewpoint is accepting the fast conclusion of the contract on the basis of an automated decision-making process when the data subject and the data controller want it and the contract is in the interests of both parties. However, this fails to recognize that, especially with regard to the consent of the data subject, the reservation of consent is a more specific provision. The requirement of the use of ADM systems for the conclusion and performance of a contract is precisely aimed at cases in which the consent or acquiescence of the data subject does not have to be given.

3.1.2 Exemption 2. National opening clauses authorizing ADM systems

A further exemption to the ADM prohibition stipulated in Art. 22 GDPR is specified in Art. 22 (2) b) for systems which are permissible under EU law or that of the various EU member states. The precondition for this is that these provisions do not infringe on the rights and freedoms of the data subject. In practice this can include national regulations which permit the use of ADM systems that are designed for surveillance purposes or to prevent tax evasion and fraud, or are supposed to safeguard the security and reliability of a specific service provider.

The new BDSG contains an exemption of this kind in § 37 (1) BDSG. The provision states that ADM systems are also permissible in cases in which the decision has been made within the context of an insurance contract, where the data subject requested an insurance benefit and the data controller has fully granted the request or partly granted it where the decision is based on the application of binding remuneration agreements for medical treatments. In the latter case § 37 (1) no. 2 BDSG confers on the data subject specific rights to express his views and to object, comparable to those specified in Art. 22 (3) GDPR (see chapter 3.1.4).

3.1.3 Exemption 3. Explicit consent of the data subject regarding an automated decision

Even if an ADM system is not required for the conclusion or performance of a contract, an automated decision can be permissible under data protection law. Art. 22 (2) c) GDPR provides for an exemption in those cases in which the data subject has given his consent to the data controller. This exemption, which will be very significant in practice, practically turns the basic prohibition of ADM systems as specified in Art. 22 (1) GDPR into a ban with consent-based permit reservation.

However, it is not entirely clear to what exactly the consent is supposed to be referring. Consent under data protection law is centrally framed by Art. 7 GDPR and refers to the data subject's declaration of consent to the processing of his personal data. However, the wording of Art. 22 (2) c) GDPR seems to refer the required consent to the automated decision itself. This is another instance where Art. 22 GDPR fails to differentiate between data processing on the one hand and the decision resulting from this processing on the other hand. Thus in this case consent should refer expressly not only to the data processing by the ADM system, but also to the circumstance of the automated decision-making. However, in the case of simple decision support systems it would be enough to give consent to the data processing only. Consent as specified in Art. 22 (2) c), when compared with the "normal" kind of consent stipulated in Art. 7 GDPR, is a specifically extended declaration. This can have systematic consequences for the information on the basis of which the data subject gives his informed consent (on consent regarding ADM systems see chapter 3.3).

3.1.4 Requirements for "exceptionally" permissible automated decisions

The numerous exemptions to the prohibition of pure ADM systems which have been described above meet specific legal requirements and limitations regarding the design of such automated decision-making systems. In cases of permitted ADM systems in which data processing is needed for entering into or the performance of a contract or those for which the consent of the data subject has been obtained, Art. 22 (3) GDPR provides for specific protective measures which the data controller is obliged to provide.

The data controller of an ADM system is principally obliged to adopt appropriate measures "to safeguard the data subject's rights and freedoms and legitimate interests." Here the regulation includes at least the right of the data subject to obtain human intervention on the part of the controller, the right to explain his or her view, and the right to contest the decision (see chapter 3.4.1). All three measures aim to make it possible for the data subject to persuade the data controller to alter a decision. However, these rights do nothing to alter lawfulness and the binding effect of the automated decision. In addition to these user rights, the data controller, within the meaning of Art. 22 (3) GDPR, can be obliged to implement supplementary measures to safeguard the rights of the user. These may also apply to the range and extent of the information-related, disclosure-related and explanatory duties (see below).

A special kind of limit to the admissibility of ADM systems is provided by Art. 22 (4) GDPR: Automated decisions may not be based on special categories of personal data as specified in Art. 9 (1) GDPR. These data include

information about racial and ethnic origins, political opinions, religious and ideological convictions, genetic or biometric data, and information about health, sex life and sexual orientation. However, this restriction on the exemptions to the ADM prohibition is once again undermined by GDPR exemptions itself. Personal data can in fact be used for automated decision-making if the data subject has given his consent, or if it is permitted as a result of a legal provision of the EU or of one of its member states.

The principal prohibition of automated decision-making specified in Art. 22 GDPR does not only have a restricted area of application. Beyond that, the regulation provides for numerous exceptions. Under the GDPR regime automated decisions will continue to be daily practice. Decision support systems, which "merely" help the decision-maker in the context of human decision-making, can be utilized to an unlimited extent within the general requirements of the GDPR. In order to elucidate the extent to which the data protection legal framework will be making a contribution from May 2018 onwards when it comes to safeguarding the normative goals one needs to take into account the possibilities of the various data protection measures, which will be scrutinized in the following chapters.

3.2 Information duties for the data controller of an ADM system *prior* to data processing

Data protection law assumes that when it comes to data protection, individual rights will be exercised primarily by data subjects themselves. The more a data subject knows about who is trying to process which personal data and for what purpose, the sooner he can actually exercise his right to informational self-determination. Transparency-related provisions, which target various data-processing entities and are manifest as a statutory duty to provide information, represent a traditional regulatory concept in data protection. In this respect, the GDPR is no different, as it also obliges a data controller to provide data subjects transparent information regarding any proposed data collection and processing. According to GDPR data protection regulations, a data controller is the natural or legal person who determines the purposes and means of processing data through an ADM system. This means that the original software developer, who may have nothing to do with how the system is applied later on, is not the data controller in the context of data protection.

The GDPR stipulates a comprehensive duty to provide information. According to Art. 13 (1) and (2) GDPR, a data controller is obliged to provide data subjects information at the time of obtainment of any personal data. This information duty includes the name of the data controller and her contact details, the purposes or changed purposes and the legal basis for data processing, the period for which personal data will be stored, the recipients or categories of recipients of the personal data (where applicable), and whether the data controller intends to transfer personal data to a recipient in a third country (also where applicable). Furthermore, before processing commences, the data subject must, among other things, be informed of his rights, like the right to request access to and rectification or erasure of personal data as well as the right to withdraw consent. In addition, data controllers are obliged to inform data subjects of the right to lodge a complaint with a supervisory authority. Art. 14 of the GDPR specifies that similar transparency duties also apply to data controllers who have obtained personal data from a source other than the data subject his- or herself.

Art. 12 (1) of the GDPR sets high standards in terms of the style and form of information communicated, which should be concise, transparent, intelligible and easily accessible, and must use clear and plain language. The data controller may use standardized icons in order to provide a meaningful overview (Art. 12 (7) GDPR). The regulation thus draws on the principles of precision and comprehensibility. The benchmark for comprehensibility, which will be difficult to put into practice, must be the "average recipient" of the information and therefore depends on the data processing context.

In terms of providing information, Art. 13 (2) f) and Art. 14 (2) g) GDPR are of particular importance for ADM systems. These articles stipulate the duties of data controllers when using ADM systems in line with Article 22 GDPR to provide a data subject with "meaningful information about the logic involved, as well as the significance

and the envisaged consequences of such processing." The term "logic involved" and the specific range of information-providing duties that derive from this provision are rather controversial (see below).

3.2.1 Duty to provide information about the "logic involved" and the "significance and consequences of data processing"

If a data controller oversees a permitted ADM system in the sense of Art. 22 GDPR, she has, in addition to the general data protection duty to provide information, a duty to provide information relevant to the specific circumstances of automated decision-making, the logic involved and the significance and envisaged consequences of the data processing. It is generally agreed that the wording of Art. 13 (2) f) GDPR specifies a duty to inform the data subject of the intent to use data processing in making an automated decision. Hence, information regarding the deployment and use of an ADM system as stipulated by Art. 22 GDPR must be supplied. The duty to provide information also exists when the data controller is not obliged to obtain the consent of the data subject but can rely on one of the other permitted legal bases, e.g., the data controller must explicitly state that he is entitled to collect and process information when this is necessary for entering into or performance of a contract. At this point, it becomes clear that the limited scope of application of Art. 22 (1) GDPR has a direct effect on the ADM-specific duties to provide information. If a decision support system which merely supports human decision-making or issues recommendations is used, it is not a case of automated decision-making in the sense of Art. 22 (1) GSPR, and thus, the data controller is not obliged to inform about the use of such a system. Moreover, the data controller does not have to provide information about ADM systems whose decisions' consequences are below the threshold of legal or real significance.

The GDPR-mandated duty to provide information regarding the use of ADM systems raises questions regarding the specific objective targeted by the regulation. When it comes to safeguards involving information about whether or not a system involves ADM, the purpose seems to be less about facilitating informational self-determination, and more about facilitating the ability to decide against becoming the object of an automated decision in the first place. Protecting human dignity appears to take precedence over safeguarding informational self-determination here. The duty to provide information about the "how" of an ADM System (i.e., the "logic involved" and the "significance and consequences of data processing" within the framework of automated decision-making) is what later on allows for the possibility of exercising the right to informational self-determination. This is particularly important for being able to assess the relevance of allowing one's own personal data to be processed by an ADM system.

The debate has focused on the duty to provide transparency in automated decision-making with regard to the "logic involved" and the "significance and consequences of data processing", and focuses on the question of the reach and depth of this transparency obligation. Some (international) legal researchers believe that this constitutes a duty to disclose the source code, whereas others interpret the stipulation as a fundamental duty to use explainable ADM systems (see chapters 3.4.2 and 7.1). Others argue that in view of the data controllers' interests in secrecy, there is merely a duty to provide abstract information about the workings and criteria of automated decision-making. All three assessments would lead to a situation in which the duty to provide GDPR-compliant information would go well beyond the former legal status in Germany. In the area of scoring (§ 34 (4) s. 1 no. 4 BDSG), a ruling by the BGH (Federal Court of Justice) has stated that the right to information applies solely to the processed data and not to the computational model that is used for scoring (BGH NJW 2014, 1235). Determining the amount of information associated with the logic involved is particularly relevant for this report, as it can result in requirements for the transparency of ADM systems as well as the individual decision which, in theory, might have a positive impact on some of the aforementioned normative objectives.

In order to determine what actually is included in the duty to provide information about ADM systems, it is useful to consider the recitals, though these are not legally binding. In Recital 71, the legislator merely clarifies the right to explanation after processing and automated decision has taken place in reference to an explanation of the ADM decision (see chapter 3.4.2). The recitals do not provide any specific insight into the extent of ex ante information (see especially Recital 60). With regard to the fact that the duty to provide information applies before the data is collected and processed, it should be pointed out that the description of the logic involved cannot refer to the result

of a single decision, and at this point in time is still focused on the automated decision-making system in its abstract entirety. The duty to provide meaningful information about the "logic" of the ADM system does not involve, at this point, the duty to ensure the explainability of or even to justify an automated single decision (see chapter 3.4.2 and 7.1).

In view of the GDPR's protective purpose, the emphasis on providing meaningful information must primarily focus on describing the data processing, its underlying concepts and mathematical models with (potential) relevance for the final decision. The data subject must be able to make an informed assessment regarding which data is (potentially) relevant to decision-making, and to what extent this can be reconciled with exercising her right to data protection and the basic rights it safeguards (e.g., religious freedom and freedom of expression). Unlike providing information about the decision-making potential inherent in an ADM system and which processing results lead to what kind of decisions, providing information on the consequences of an automated decision itself are not a central aspect of the duty to provide information prior to data collection.

The range of this principally abstract description of a system and of its decisions' underlying assumptions, assessments and concepts is restricted by two partly contradictory facts: On the one hand, the specific nature of the information is limited by the controller's interest in secrecy. Where a description of the system reveals too many details of the data processing and decision-making involved, the duty to provide information, which was originally enacted as part of the right to data protection, runs up against similar legally protected rights and freedoms of the data controller and/or of the system developer (i.e., the freedom to choose an occupation, the freedom to conduct a business and the right to property; see Art. 15, 16, 17 EU Charter of Fundamental Rights). On the other hand, the duty to provide information is determined by the legal requirements regarding this abstract description. A mathematical description of how a system functions will strike the average user as being precise, but not easy to understand, and thus not particularly meaningful. The valid duty to use comprehensible language, which also applies to the comprehensibility of the "logic involved," also limits the concrete mathematical and technical description of the ADM system – irrespective of whether such a description is even theoretically possible, given the architecture of a system (on the explainability of machine learning and artificial intelligence-based systems see chapter 7.1). Nevertheless, a description that is too abstract and does not reveal enough details gets into danger of not complying with the duty to provide "meaningful" information.

These contradictory requirements force the data controller to steer a middle course with regard to the depth of information, the degree of detail and the length of description regarding the logic involved. When evaluating whether or not legal requirements have been met, supervisory authorities and the courts therefore have a margin of appreciation in terms of the assumptions made. This includes assumptions about the average recipient, his cognitive abilities and an assessment of requirements regarding the extent, form and presentation of information.

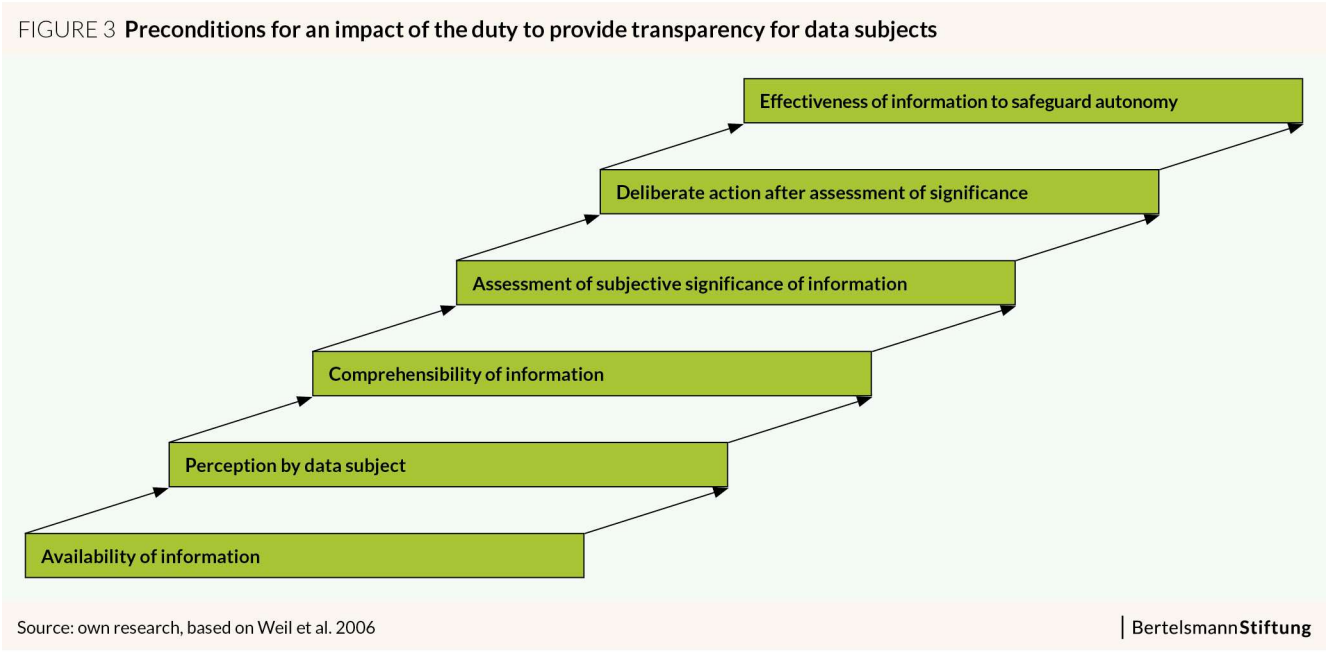
In Germany, § 32 (1) and § 33 (1) no. 1 BDSG restrict the duty of public authorities to provide information within the meaning of Art. 13 and 14 GDPR if this otherwise would "threaten the due performance of [...] tasks within the meaning of Article 23 (1) a) to e) GDPR or endanger public order and security or in some other way present disadvantages to the welfare of the Federal Republic or one of the Federal states." While legal exemptions to the GDPR's duty to provide information are specified under Art. 23 (1) GDPR, the reservations specified in BDSG do not seem to align with this, or are at least not specific enough. If data subject does not receive information, § 32 (2) and § 33 (2) BDSG specify a duty to provide specific obligations. In view of these requirements, the authorities should provide – publicly and in advance – information about the fact that they are using an automated decision-making system (Martini/Nink 2017).

3.2.2 Impact assumptions and limits of transparency provisions in the case of ADM systems

The duty to provide information is a classic regulatory measure in data protection law. The direct protective purpose of data protection law is to guarantee the right to data protection – a right that focuses on the autonomy of individual action and decision-making in light of the data associated with the individual. Individuals should be able to make decision regarding who knows what and when about them. Thus, in legal terms, safeguarding this right aims to facilitate *informed* decisions.

The intended mechanism of informational regulation that is required for safeguarding autonomous decisions is rather complex (see Fig.3, after Weil et al. 2006). First, the information needed for an individual's decision-making has to be made available. Second, the data subject must become aware of its existence. This stage of being made aware is followed by the cognitive process of understanding the content of the message. However, understanding is not enough: In order to recognize the relevance of such a message, the data subject must first integrate this information into her own system of norms and values. Only if an individual deems an information to be of actionable significance he or she can act in response.

Fig. 3: Preconditions for an impact of the duty to provide transparency for data subjects



Whenever transparency obligations, after mastering these complex steps, aims at better (i.e., more autonomous and informed) decisions by end users and end consumers, the legal system implies that the data subject is always able to act and decide rationally. However, as behavioral economics have shown, this kind of level-headed decision-making is increasingly taking a back seat to other forms of thinking (Howells 2005). Whenever information processing capacities are exhausted, attention spans are overtaxed, attention is distracted or the cognitive ability to anticipate the consequences of action are restricted, or where oligopolies, powerful network or lock-in effects, not to mention emotional pressure or an objective lack of economic motivation curtail decision-making autonomy, exercising the right to informational self-determination can be significantly restricted in terms of rationality. Given this state of research, transparency provisions amount to a theoretical regulatory approach which, in practice, does not necessarily safeguard the autonomous exercise of a right (cf. Edwards/Veale 2017: "transparency fallacy"). The duty to provide information as a tool safeguarding basic rights has limited impact in everyday life. This would change, however, if the duty to provide information on the logic involved were construed to require the use of systems able to guarantee a certain degree of explainability. In such cases, the GDPR provisions would, in addition to safeguarding transparency, have a direct impact on the comprehensibility of system decisions.

All in all, the GDPR, with the duty to provide information as specified in Art. 13 (2) f) and Art. 14 (2) g), provides measures that make transparent the implemented models, which, in theory, can have positive effects when it comes to safeguarding individual freedoms. This way, transparency might safeguard non-discrimination by revealing and abolishing discriminatory criteria. However, comprehensibility requirements place limits on the scope and depth of the information to be made available, which precludes an external evaluation from taking place that is otherwise needed if the desired individual, group and societal objectives are to be met. Because the scope and depth of the required information must be understandable in layperson terms, it is not suitable for an external ADM evaluation conducted by experts. Such audits are needed, however, in order to identify the risks for individual, group and societal objectives.

3.3 Conditions for user consent in the case of automated decisions

Conditions of user consent under the GDPR can involve a duty to inform which can reinforce or expand existing transparency duties in cases where data processing relies on another other legal basis than consent. This way, the GDPR can have positive effects for threatened rights and objectives. Reservation of consent can, above all, protect personal autonomy and personality rights, since data processing and its consequences derive from a user's informed and conscious decision; i.e., the data subject can decide for himself whether he wishes to subject himself to automated decision-making and the required data processing.

As explained above, there is a need for ADM-specific consent whenever a system of purely automated decisions has legal or other significant consequences for the data subject and there is no other legal basis for data processing (necessity for entering into or performance of a contract; legal permission). In practice, consent is a key legal solution to the use of ADM systems. Once consent has been given, it is even possible to include sensitive personal data, as specified by Art. 9 GDPR, in data processing. However, a distinction needs to be made between data subject's consent to the data processing in general (Art. 6 and 7 GDPR) and his or her specific consent to data processing within the context of an ADM process (Art. 22 (2) c) GDPR).

With regard to valid consent, the GDPR lists a number of conditions that have to be met cumulatively in order for a data controller to make use of the exemption stipulated in Art. 22 (2) c) GDPR. In addition to the duty to provide documentation of each consent in terms of Art. 7 (1) GDPR, the data controller has the duty to provide transparent information with regard to the content and the form of consent. Art. 4 no. 11 GDPR stipulates that consent must be given in form of a freely given, specific, informed and unambiguous indication of the data subject's wishes. A precise description of the purpose of the proposed data processing is paramount here. The request for consent, especially in context with other declarations, has to be highlighted and made apparent so that the can give her (electronic) consent in a conscious and unambiguous way. It is thus up to the data controller to create a situation that results in the user's awareness that she is giving her informed consent. Before declaring consent, a data subject should be made aware of her right to withdraw consent (Art. 7 (3) GDPR). There are additional requirements relating to consent which involve the processing of special categories of personal data and consent given by minors under the age of 16 (Art. 8 and 9 GDPR).

The specific consent-related requirements establish a transparency obligation for ADM systems that is coherent with the information obligations stipulated by Art. 13 (1) and (2) of the GDPR. The information specified in Art. 13 (2) f) concerning the existence of an ADM system has to be part of of the consent's scope. Otherwise, it would be impossible to speak of an informed consent. The same applies in cases where a data controller changes his data processing system into a purely automated decision-making system. Here, again, the specific ADM-related consent of all users has to be beseeched. If a data subject's consent is required in the context of an ADM system, its informed nature must also extend to the explanations of the logic involved as well as the significance and consequences of its data processing. The required level of informed consent is reached only if the data subject, before giving his consent, can assess which data will be processed by the system and in what form, and can appreciate the significance of his consent. The requirements relating to the appropriate nature of the measures designed to safeguard the rights of the data subject (cf. Art. 22 (3) GDPR) also support this view.

In addition, Art. 22 (2) c) GDPR allows for automated individual decision making in cases where explicit consent is provided. It is not clear, though, to what extent this provision constitutes an explicit obligation to inform about the existence of an ADM system within the context of a request for consent to data processing as stipulated by Art. 7 (2) GDPR, or whether a specific consent regarding the use of an automated decision-making system must be given separately. The limits of transparency as a resource to safeguard individual rights and freedoms apply here, too. The intention to use an approach that is based on free and informed consent is reasonable, but it can be weakened by the specific decision-making situation, as shown above.

The right of an individual to withdraw his or her consent as stipulated in Art. 7 (3) GDPR also applies to the consent given to data processing by an ADM system. Scholars have to clarify, though, whether and how such a withdrawal is possible if it is declared after data has been collected but before any automated decision-making has begun.

All in all, the reservation of consent as an exception for prohibited ADM systems seems to be suitable for safeguarding the individual rights and freedoms of data subjects in general, such as autonomy and personality rights in particular. But both the possible context-related weakening of the freedom to choose (e.g., through the non-existent alternatives of non-automated decision-making systems) and the limitations to the information that must be made available within the GDPR framework and its limited protective effects (see chapter 3.2.2.) undermine the effectiveness of this protective measure. Indeed, building on consent as a key privacy measure in safeguarding individuals' rights has been fundamentally criticized by some legal scholars (Radlanski 2016).

3.4 Rights of data subjects after data processing by ADM systems

In addition to those tools that safeguard normative objectives prior to data processing, the GDPR also provides for the protection of a data subject's rights and freedoms by measures addressing the point in time after data is first processed by an ADM system. At this juncture, the GDPR assigns to the data subject a whole series of legal rights. In addition to the right to access personal data (Art. 15 GDPR), we must mention the right to object (Art. 21 GDPR), the right to rectification (Art. 16 GDPR), the right to restriction (Art. 18 GDPR), the right to erasure (Art. 17 GDPR), the right to data portability (Art. 20 GDPR), the right to lodge a complaint with a supervisory authority (Art. 77 GDPR), and the right to compensation (Art. 82 GDPR). In addition to these general data protection rights, the regulation also contains ADM-specific provisions.

3.4.1 Right to obtain human intervention

Art. 22 (3) GDPR obliges the data controller of an ADM system to "implement suitable measures in safeguarding the data subject's rights […] 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." Despite the wording, it is assumed that Art. 22 (3) GDPR does not give the individual a subjective and enforceable right but instead obliges the data controller to organize the decision-making process in a certain way. The obligation to provide a "human in the loop," who, if the user makes an appropriate request, is called upon by the data controller to participate in the process, mirrors and specifies the fundamental right of the individual not to be subject to automated decision-making as specified in Art. 22 (1) GDPR. The reference to the dignity of human beings is once again apparent here (Zarsky 2017). This provision empowers the data subject to obtain – ex post – human intervention in scrutinizing an initially automated decision including the underlying processing of personal data. It also empowers a data subject to restore human decision-making after clarifying the data subject's own perspective. If the potential for human intervention as specified in Art. 22 (3) is to protect the data subject's rights, at least two preconditions must have been met. On the one hand, the human called in to intervene must have respective leeways in decision-making in order to make his own assessment and take a new decision on the basis of the facts of the case - which may well have been modified by ex-post remarks made by the data subject – including to revise the automated decision in this single case. On the other hand, in order to make his own assessment, the appointed decision-maker – particularly in complex cases – must know which factors have led to an automated decision, as these are required to facilitate a new evaluation.

Scholars have claimed that the right to obtain human intervention is not available to every data subject without exception, for this would be in systematic contradiction to the GDPR, which in Art. 22 (3) GDPR makes requirements on ADM systems which within the wording of Art. 22 (2) are exceptionally permissible, where the general principle bans pure ADM systems in Art. 22 (1) GDPR. If everyone were entitled to the right of obtaining human intervention, the principle of general prohibition and exceptionally permission would be nullified, because it would theoretically be possible to call for human intervention after every single decision. As a result, the principle of the impermissibility of automated decisions would apply without exception. Against this background, the right to human intervention may be restricted to legitimate individual cases – "the minimum protection of personality rights in automated processes includes the possibility that the data subject can point at matters that show the unique relevance of an

individual case" (Martini/Nink 2017). However, this restriction means that the data subject will always have to justify human intervention, which will again result in a structural problem. As the controller's duty to provide information in a comprehensible way is limiting deeper insight into the decision-making process, this, in turn, makes it difficult for the data subject to identify a reason for such a justified application. The only way forward would be to exercise the right to lodge a complaint with the appropriate supervisory authority. However, one could also interprete this restriction to justified cases as a requirement for the notions of comprehensibility. In this case, the dtaa controller might have to create an informational situation in which a neutral observer could decide whether a case is an unusual one which exceptionally justifies the demands for obtaining human intervention.

With the right to obtain human intervention the GDPR provides a legal tool that makes it possible for a human being to verify an individual case of automated decision-making. The assessment of an ADM decision by a human included in the process is a relevant means of safeguarding individual rights, as this person can verify the facts, including additional statements by the data subject, and has the power to take an amended decision. However, the right to obtain human intervention merely means that one is entitled to a repeated (human) decision – not to one which the data subject might consider to be an improvement. The possibility of verification does not necessarily lead to a result that differs from the former unsatisfactory automated decision. Furthermore, given the lack of precedents in applying the right to obtain human intervention, the scope of its application cannot be predicted yet. A limited interpretation of this right may mean that it applies only to a fraction of the ADM processes in practical use.

3.4.2 Ex-post right to information about automated decision-making

Another right that specifically targets ex post data processing in ADM systems is the right to information as specified in Art. 15 (1) of the GDPR which, under letter h), contains a provision with wording similar to that found in the ADM-specific duty to provide information before the data processing occurs (see chapter 3.2.1). Besides general information on the processing purpose, data categories, envisaged periods of storage and on user rights, the user's right to ex post information regarding automated decisions includes specific statements on the logic involved and an explanation of the significance and the envisaged consequences of data processing. Notably, the right to information is an individual right which the data subject can exercise after the data have been processed. This can have consequences for the volume, the source of, and the way in which the information is provided. The right to information includes information about measures being taken to meet the requirements specified in Art. 22 (3) GDPR inasmuch as these are appropriate measures designed to protect the individual's freedoms (see chapter 3.1.4). This can include internal procedural measures which, for example, are concerned with the monitoring, evaluation and optimization of automated decision-making processes (see chapter 3.5).

There are problems associated with the right to information, though, that derive from the possibility that a data subject may exercise his or her right. A consequence of this right is that the data controller is compelled to store all decision-making data and to ensure their availability in the event of a request for information – even in cases where there are no other reasons to do so. This undermines the principle of data minimisation. Legal scholars have thus suggested that data controllers should be allowed to limit such storage in material and temporal terms. After a certain period of time has passed, the right to information then would, in practice, be impossible or virtually impossible to enforce.

Moreover, scholars disagree over what substantially transpires from the right to information in terms of duties as specified in Art. 15 (1) h) GDPR. The perceived extent of the duty to provide information in the case of automated decisions ranges from a duty to provide information that is restricted to the processed data and the data processing formalities similar to the duty to provide information as specified in Art. 13 and 14 (see chapter 3.2; Gola 2017), via criteria-based record-keeping (Ernst 2017) or the disclosure of the automated results (Bäcker 2017) to the decisions' reasoning (Goodman/Flaxman 2016). Sometimes, the legal framework is taken to mean that it is permissible to use only ADM systems which are explainable ("legible systems;" Gianclaudio/Comandé 2017). The question of the extent to which the specific requirements of the right to access generate a "right to an explanation" of the automated decision has been the subject of international debates (Selbst/Powles 2017). One opinion commenting on transparency provisions distinguishes between the ex-ante information about system-related

concepts and models (as required within the framework of Art. 13, 14 GDPR) and the required ex post explanations within the framework of access rights (as specified in Art. 15 GDPR), which focus on single cases and thus comprise the reduction of abstract explanations to single-case evaluations or assessments (Wachter et al. 2017). Here the authors talk of a "limited right to explanation." These claims that such a right to explanation is specified in Art. 15 (1) h) GDPR – here, the opinions remain rather implicit – are tightly associated with potential threats. Such demands for the explainability of automated decisions are driven by the understandable desire for better forms of "accountability" among data controllers of ADM systems that would allow for identifying and criticising flawed decision-making systems (regarding safeguarding "accountability" see chapter 7.2). However, when interpreted in this manner, such demands are also mainly based on the improvability of decisions and their consequences for individual, group or societal interests in terms of preventing discrimination.

To date, the academic community has not come up with a differentiated analysis of the protective purposes of the GDPR. As explained, data protection is primarily concerned with data-processing related protective approaches as its preliminary protective. However, data protection law also protects downstream basic rights and freedoms (see chapter 3). The risks posed by automated decision-making – if one ignores the necessary data processing that it involves – do not impinge directly on data protection safeguards, but usually on aspects of human rights connected to personal autonomy and human dignity, inter alia. This is the case especially when subjects in automated processes become objects of mathematical and probabilistic calculations. Not only data protection goals such as the protection of the right to privacy seem to be threatened by automated decisions and their consequences; but so do goals such as freedom from discrimination as well as group-related and societal goals such as social inclusion and fairness. But whether and to what extent a right to explanation (discerned from Art. 15 (1) h) of the GDPR) that is based on a complaint lodged by a data subject is in any way suited to safeguard these protective purposes, has not been in the focus of the debate. Yet this question is relevant, given the provisions regarding the style and content of the obliged information described above, which apply to the data controller whenever she provides information to the user. The duty to provide the user with comprehensible information is bound to lead to limitations with regard to the depth and detail of the transparency briefing. As a result, the user's right to information in particular is too limited to enable him, for example, to identify systematic bias, erroneous conceptual assumptions, or faulty weightings of decision-making factors resulting in discrimination of individuals or groups – especially since it's not possible to draw conclusions about systemic decision-making through the explanation of an individual case.

This differentiated interpretation (especially with regard to reaching the legislative purpose) of the ADM-specific information duty shows no basis for an in-depth explanation of ADM decisions under Art. 15 (1) h) GDPR. However, the data subject's right to information contains an ex post explanation of the underlying decision-making model and system functionality in terms of processed data categories, including a depiction of how individual factors are weighted in a particular case (see also Article 29 Working Party 2017). In practice, this may make it possible to see the reason, for example, for a negative decision, though the data controller is not compelled to justify his weighting choice nor the decision in a rational sense, but merely to provide information about the automated decision process and its underlying derivations. The right to information is geared to information about the decision-making process, not to the interpretation or legal explanation of its results.

Thus, the right to information helps – above all – to safeguard threatened individual rights and freedoms such as personality rights or an interest in fair treatment by ADM systems. In-depth insights into the ADM system – which might allow for identifying possible decision-making bias and resulting discrimination of specific societal groups on a structural level – will not be provided by the explanation of one particular case.

3.4.3 Excursion: Justification duties in cases of administrative decisions

In Germany, there are specific public law provisions that apply to automated decisions of public authorities. Here, for constitutional rather than data protection reasons, legal stipulations prescribe a duty to justify public decisions. In § 35a of the Verwaltungsverfahrensgesetz [Administration procedure act] (VwVfg), which came into force on 1.1.2017, the legislator allows fully automated administrative decisions only "if this is permitted by a legal provision" and the public authority concerned has "neither discretionary powers nor margins of appreciation." For the latter,

an automated process does not come into question. Moreover, every automated administrative process requires a specific legal provision allowing a public authority to use ADM systems.

But the procedural provisions for the justification of administrative acts as stipulated by § 39 VwVfG also apply to automated decisions. As a result, essentially every written or electronically issues administrative act is to be accompanied by an explanation which contains the relevant "material and legal grounds" for the decision. The provision of a justification can be waived if an application has granted an application and the administrative act does not impinge on the rights of third parties, or if the authority "issues identical administrative acts in considerable numbers or with the help of automatic equipment and individual cases do not merit a statement of grounds". Thus, automated processes can be used in cases of simple cases and corresponding administrative acts, which do not require an explanation. On the other hand, administrative procedures in which the competent authority is granted discretionary power or a margin of appreciation cannot be automated. No justification is necessary only in the case of procedures involving areas as specified by § 35a VwVfG and which do not require an explanation. The legal analysis of the compatibility of administrative law provisions relating to automated decision-making with the new GDPR requirements continues to be difficult, even in these cases (Martini/Nink 2017).

The VwVfG does not stipulate a duty to provide an explanation where automated decisions are advantageous from a constitutional point of view and in straightforward, simple cases. However, for all other ADM systems used by administration, the principle of the duty to justify applies and goes a step further than the data protection-based explanation. Constitutional principles require a legally compliant explanation of the automated decision or wholly prohibit the use of ADM systems in cases where administrative leeway in decision-making is exercised. Guaranteeing individual rights and freedoms in administrative matters appears to have been afforded high priority.

3.5 Systemic and procedural duties of providers of ADM systems

In addition to the information obligations towards the user as well as the requirements relating to the implementation of automated decision-making processes that emanate from user rights, GDPR provisions also contain obligations related to the process or the design of the system, which might have direct or indirect supportive effects regarding the achievements of the goals described above. Thus, general qualitative requirements with regard to the use of certain automated calculation or decision-making processes, but also process-related provisions for the design and implementation phases can help guarantee the interests and rights of individuals, groups and possibly even society as a whole.

3.5.1 Compulsory use of approved mathematical methods?

A minimum level of quality assurance in the design of automated decision-making systems can be achieved by compelling the data controller to use approved mathematical methods. Such a provision is specified in § 31 BDSG n.F. for scoring procedures, i.e. for a narrower area of application than Art. 22 (1) GDPR. The requirements of exceptionally permissible ADM systems (Art. 22 para. 3 of the GDPR) do not include such an explicit restriction on approved methods. However, in Recital 71 the legislator explains that data controllers should "use suitable mathematical or statistical methods for profiling." Considering the non-binding nature of the recitals and the use of the rather weak word "should", only the interpretation of other provisions of the GDPR might result in a legal requirement to use approved mathematical methods.

There are a number of possibilities, including the general duty to adopt appropriate measures to safeguard the rights of data subjects when using ADM stipulations as specified in Art. 22 (3) of the GDPR as well as the general principle that the data controller will assume responsibility as defined in Art. 24 (1) of the GDPR, taking into account the system's state-of-the-art technology on the basis of the "privacy by design" principle and possibly the requirements with regard to data security specified in Art. 32 (1) of the GDPR. All four norms oblige the data controller to take appropriate measures to defend the rights, freedoms and legitimate interests of the data subject. However, this does not fully exclude the use of (hitherto) uncommon mathematical methods. As long as the provisions state – and they do so explicitly in Art. 25 (1) and Art. 32 (1) of the GDPR – that the data controller should

take his bearings from the state of the art, his restriction to use only approved mathematical methods can be a clear indicator for. The use of recognized mathematical methods can thus be of importance when examining the rightfulness of these measures. However, there is no explicit obligation to choose only approved methods as long as the provider is merely obliged to take his bearings from the state of the art and can decide to use his own methods, which may be just as good. An interpretation as restrictive as this would risk encouraging the growth of algorithm or ADM system monocultures. The risk of a greatly reduced level of system-related diversity alluded to above would be alarming. However, the duty to take one's bearings from state-of-the-art technology can basically rule out the use of ADM processes that are completely untested in the wild and bear the risk of being faulty.

3.5.2 "Privacy by design"

The GDPR contains specific requirements relating to the technical design of data processing systems (Art. 25 (1) of the GDPR). The regulatory approach known as "privacy by design" obliges the data controller in the interest of the user's best protection at the earliest opportunity – for example during the product development – as well as during and after the implementation of the system, to consider data protection issues in an ongoing and systematic manner. Within this concept, basic principles are data avoidance and data processing minimization, as well as data anonymization and pseudonymization. The data controller is explicitly called upon to implement organizational measures besides technical ones: in addition to system-related arrangements there should be organizational measures such as training courses or interdisciplinary project groups. The size and scope of the required measures are dependent on the nature, extent and purpose of the data processing, the associated risks for the data subject and the cost of the necessary implementation. Non-compliance with these provisions – in contrast to the statements contained in § 3a of the old BDSG – can lead to the imposition of a fine.

The threats to the freedoms and rights of the individual emanating from the use of automated decision-making processes have to be taken into account in a systematic manner when it comes to conceptualize, develop and ADM systems. They should also be prevented from becoming a reality with the help of such a risk management approach and appropriate counter-measures. However, group and societal risks are not covered by Art. 25 (1) of the GDPR.

3.5.3 Additional obligations with potential positive impact

A series of additional provisions in the GDPR – comparable to the duty to use data protection-friendly technology and organizational design as specified in Art. 25 (1) of the GDPR – can also ensure that the data controller is duly aware of data protection risks. They include Art. 30 of the GDPR, which is of central importance and specifies a duty to compile records of all processing activities, as well as Art. 35 (3) a) GDPR, which specifies a duty to carry out a data protection impact assessment (DPIA) specifically for ADM systems. The two provisions taken together, along with Art. 25 (1) GDPR, provide an early opportunity to consider the minutiae of the planned data processing, including whether or not the planned implementation is in line with the GDPR provisions. Whereas privacy by design analyzes and evaluates on the basis of risk management – i.e. minimizes risks but does not fully exclude them – a DPIA seeks to identify and put a stop to any infringement of the GDPR from the user's point of view. The inspection within the framework of the DPIA covers not only the technical systems needed for data processing and storage, but also the organizational arrangements made by the data controller. In general, Art. 35 GDPR leads to a systematic ex ante self-evaluation of the technology and processes used by ADM systems. But like the GDPR, the DPIA is limited to the rights, freedoms and justified interests of the individual.

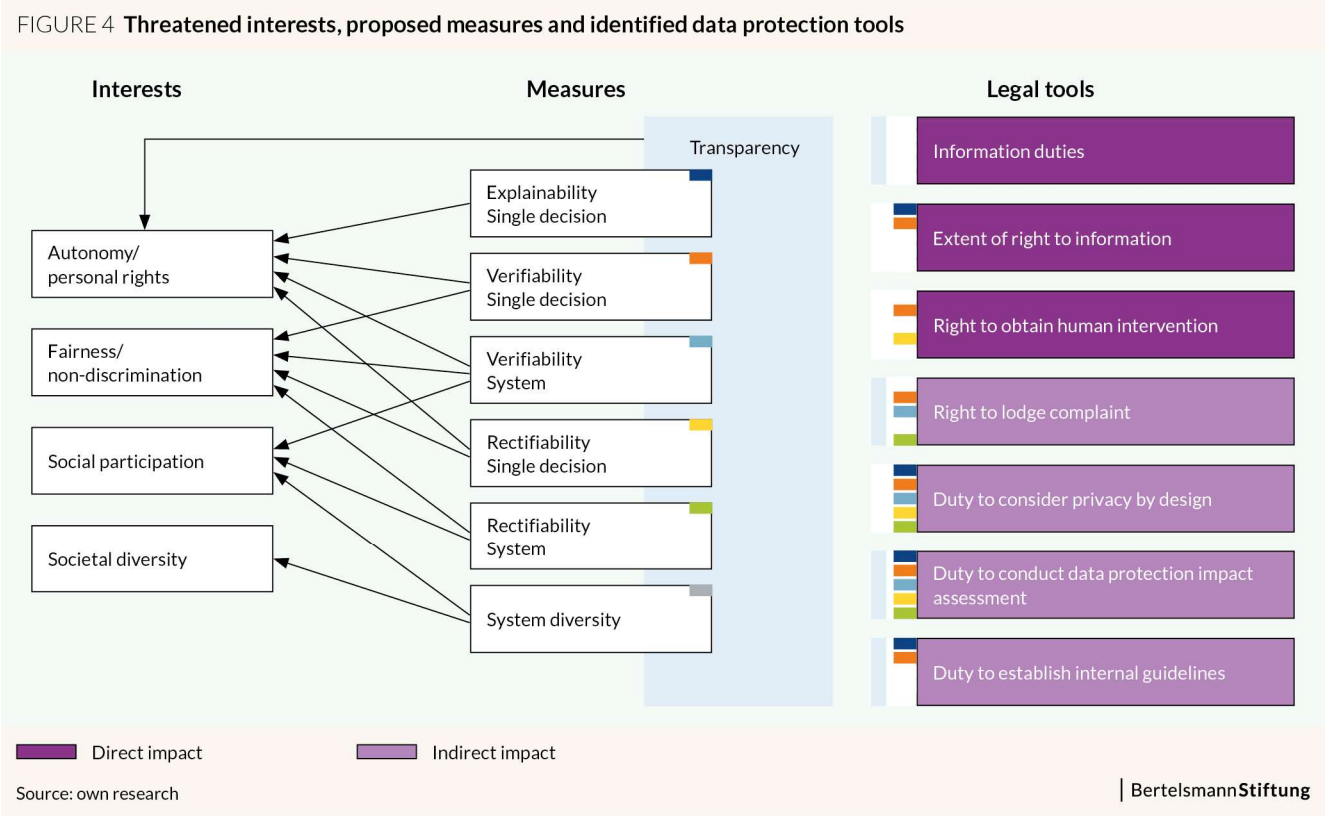
These duties are flanked by Art. 47 GDPR, which approves binding corporate rules with regard to ADM systems (Art. 47 (2) lit. e). These rules must contain regulations concerning the rights of data subjects and thus promote a high degree of awareness regarding data protection on the part of the data controller. The appointment of a data protection officer can help to secure the systematic inclusion of the data protection perspective in the planning, implementation and internal evaluation. A compulsory appointment in case of ADM systems provision is not specified in Art. 37 GDPR. However, the special provisions in § 38 (1) BDSG n.F. stipulate that data controllers must appoint a data protection officer if – according to Art. 35 of the GDPR – they are obliged to carry out a data protection impact assessment. As explained above, this is indeed the case, which creates a duty to appoint a data protection officer at least for providers of automated decision-making processes with a branch office in Germany.

4 Where the GDPR brings benefits: Approaches to safeguarding individual rights and freedoms

After the analysis of the goals endangered by ADM systems and possible counter-measures, we can map the legal provisions in the GDPR and BDSG n.F. applicable to automated decisions regarding their potential to protect freedoms and public interests. Areas that lack protective tools are areas of risk where weak legal frameworks currently provide no counter-measures to protect the goals alluded to above. With regard to this mapping, it is important to keep in mind that these protective effects are only hypothetical for now. Neither the GDPR nor the BDSG n.F. had come into force at the time of this analysis; many of the undefined legal concepts and norm interpretations are the subject of rather heated academic and political debate. Controversy should lessen once they have entered into force, as data protection routine, the supervisory authorities and normal court rulings will create a more or less legally secure status quo in the upcoming years.

In order to clarify the question of the extent to which the data protection legal framework can contribute to safeguarding normative goals from May 2018 onwards, goals and the measures adopted to safeguard them are compared with the data protection tools of the legal framework for ADM systems (Fig. 4).

Fig. 4: Threatened interests, proposed measures and identified data protection tools



The legal analysis has shown that the legal framework consisting of the GDPR and the new BDSG has at its disposal tools that can have positive effects on one or more of the stipulated goals. While some of these tools have a direct impact via information obligations, user rights and petitions, others work rather indirectly where data controllers are obliged to adopt precautionary measures and to conduct ex ante evaluations and assessments.

4.1 Transparency rules strengthen autonomy and personality rights of the individual user

Information duties and users' access rights are positioned to create transparency for the data subject. Thus, the data subject's knowledge about the existence of an ADM system in general as well as its proposed use, the nature of the data collected and the purpose of the data processing can help the data subject assess the processing and the decision-making in terms of their relevance for his or her autonomy and personality rights. In cases of ADM systems in which the controller requires the consent of the user, requiring an explicit consent regarding the use of an automated decision-making system improves the user's ability to exercise personality rights. However, both levels still show limited effectiveness regarding these forms of transparency as the basis for conscious and rational decisions by the user: Transparency does not necessarily lead to an optimization of basic rights. While it may improve the decision-making basis of the data subject, it is all but certain to do so in practice (see above chapter 3.2.2). Furthermore, the duty to provide information corresponds to the context-dependent average user's level of comprehension. Thus, the extent and depth of information stemming from the GDPR's information duties are specifically constrained by the requirements pertaining to the comprehensibility of the provided information.

The right to information, focusing on single decisions and meaningful disclosure regarding the logic involved go a step further, as the data controller must also explain the factors comprising the decision and its significance for the outcome of the decision-making, although here again in a comprehensible manner for the user who has exercised the right to information (see above chapter 3.4.2). For single cases, rights to information at least enable the data subject to assess in approximate terms whether an automated decision is reasonable and valid; in cases where discriminatory treatment is suspected, the data subject can exercise further rights then. Thus, disclosure rights relating to ADM systems function as a starting point for safeguarding personality rights and as a guarantee of fairness and social inclusion for the individual user.

4.2 Rights to obtain human intervention guarantee a "human in the loop"

In the case of a purportedly wrong decision or what seems to be unfair treatment, the data controller must provide procedures that enable the data subject to obtain a human intervention. If properly addressed by the data controller, such intervention will lead to a human reassessment of an automated decision and the possibility of a revised decision on the part of the provider, also taking into account unusual facts and special circumstances that were not (or could not be) included by the ADM system. As shown above, one interpretation of the GDPR proposes a minimum requirement such that the data subject is able to deduce from the explanation whether this is a special case that entitles the data subject to human intervention. This important right to obtain human intervention might help in facilitating the verification of a single decision made by the system, but it is intended above all to enable a new, human-based decision. Thus, the right to obtain human intervention appears to be the principal legal tool for verifying and rectifying an automated decision ex post facto. Both verification and rectification underpin the safeguarding of personality rights and fairness as normative goals.

4.3 Positive indirect effects of systemic and procedural duties

The encompassing duties of ADM system providers to consider privacy-by-design principles, to carry out data protection impact assessments, to compile internal guidelines and to appoint a data protection officer encourage a high level of awareness with regard to the potential risks of their ADM processes. This leads to a strong consciousness regarding the rights of data subjects, already in the early design stage as well as after the implementation of the system, establishing a rights-aware framework regarding all ongoing organizational processes. The appointment of a data protection officer results in a systematic position in the organization focusing on data protection issues. In general, the systemic and procedural regulations are of great importance with regard to normative goals: During the design and implementation phase of ADM systems comprehensive evaluation and

assessment duties designed to safeguard user rights can help data controllers detect and minimize possible risks for individuals – sometimes as well as for groups of people (see chapter 5) – at an early stage. With regard to the ongoing control of the system – including addressing cases in which users have challenged an automated decision or exercised their right to obtain human intervention and have thus identified a potential fault – the respective legal duties constitute a strong incentive to introduce simple procedures for the examination and rectification of individual decisions and to monitor the system once it is up and running. And last but not least, the data subject always has the right to lodge a complaint, in which case the competent data protection supervisory authority is called in. The obligation to "implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests" (Art. 22 (3) GDPR) in the design and implementation of ADM processes focuses on individual user rights and freedoms. It also makes it possible for the data controller – in case faults are actually detected – to adopt internal measures such as the verification and rectification of single decisions and of whole processes, which contributes to the safeguarding of personality rights, fairness and social inclusion, especially with a focus on the individual.

As the systemic and procedural duties result in internal measures of the data controller only, the actual implementation needs an effective incentive system. Here, the GDPR and the new BDSG use the traditional concept of institutionalized administrative and supervisory bodies: the data protection commissioners on the Länder and the federal level as competent institutionalized supervisory authorities (see chapter 4.4).

4.4 Role and measures of data protection authorities to safeguard the normative goals

The supervision of compliance with the legal provisions of the GDPR and the new BDSG falls to the data protection authorities. Their central tasks include the supervision and enforcement of the GDPR (Art. 57 (1) a) GDPR). For this purpose, the supervisory bodies have been given far-reaching inspection powers and access rights (Art. 58 (1) GDPR). These include, inter alia, the power

- "to order the controller and the processor, and, where applicable, the controller's or the processor's representative to provide any information it requires for the performance of its tasks" (lit. a));
- "to carry out investigations in the form of data protection audits" (lit. b));
- "to obtain, from the controller and the processor, access to all personal data and to all information necessary for the performance of its tasks" (lit. e));
- "to obtain access to any premises of the controller and the processor, including to any data processing equipment and means (…)" (lit. f)).

The combination of these powers clearly demonstrates that the inspection and information rights of the authorities – in contrast to those of the data subjects – are not limited by requirements such as comprehensibility or by being restricted to one particular case. The supervisory authorities' criteria for any examination derive from their task, which primarily is to supervise and enforce the implementation of GDPR norms. As shown above, the focus of the GDPR is on the right to data protection and on other basic rights and freedoms of the individual. Thus, the supervisory authorities also focus specifically on potential infringements of or risks to individual legal rights (but see below, chapter 5). Within their competences, the supervisory authorities are empowered to inspect any data processing installation, including the software being used as well as the connected databases. This way, they are able to audit compliance with the GDPR regulations in close proximity to the automated decision-making processes. The data controller has an encompassing duty to provide information to the data protection authority whenever such inside knowledge is required. Theoretically, the GDPR's governance structure is positioned to provide powerful incentives for the implementation of the systemic and process-related duties of the data controller. If addressees of the regulations are aware that all concepts, impact assessments and software programs relating to their ADM systems can be inspected by the data protection authorities during a data protection audit, the data controller's motivation to carry out such project phases with great care is likely to increase. Thus, the governance

structure envisaged by the GDPR and the new BDSG is able to ensure the implementation of the systemic and process-related provisions, and thus to support the legal objectives alluded to above. In order to let these structural considerations become reality, it will above all be necessary to exert a certain amount of supervision and enforcement pressure. In light of the fact that some of the data protection authorities have already reached their capacity limits, the extent to which this pressure will be built up remains to be seen.

4.5 Interim conclusions: GDPR's starting points for measures safeguarding individual rights and freedoms

In essence, data protection law protects the rights and freedoms of the individual. It protects directly and without reservation the autonomy and personality rights – in particular the right to privacy – and, indirectly, the exercise of other basic rights and freedoms threatened by data processing. The legal framework aiming at the protection of individual rights is positioned to create useful transparency for the data subject concerning the functioning and underlying models of ADM systems in all single cases. Furthermore, the explicit information about the fact that an automated decision on the basis of personal data will take place allows the individual person to take a respective decision to protect his or her autonomy. However, the amount and extent of the disclosed information are limited by the requirement to keep a user-friendly level of comprehensibility. Moreover, the duty to provide information for the end user has a limited effect only in terms of the (rational) perception of an individual's autonomy. The right to obtain human intervention, which is framed by the rights to express a point of view and to lodge a complaint, and the resulting possibility, facilitated by the provider, of a subsequent review and new decision are positive in terms of the rectifiability of automated decisions. However, the data subject is not empowered to more profoundly understand the underlying assumptions and concepts. Thus, the data subject's rights as specified in the GDPR facilitate – at least partially – the attainment of normative goals such as autonomy of action and the protection of personality rights and can rectify cases of individual discrimination.

Above that, the GDPR requirements regarding systemic and procedural organizational requirements – underpinned by the far-reaching supervisory and monitoring powers of the data protection authorities – can compel the providers of automated decision-making processes to take a systematic interest in the issue and to develop a high level of awareness of the possible legal threats to individuals. Careful thinking and powerful incentives for the establishment of risk-centered processes can create a situation in which, during the design and implementation phases, the data controller treats automated decisions responsibly and procedures for verification and rectification are firmly in place. However, a negative aspect of these duties remains: the provision of such safeguards occurs solely within the context of internal and non-public deliberations and development procedures. If there is a positive effect, it will always remain indirect.

5 Where the GDPR falls short: Weak spots concerning group-related and societal interests

The GDPR's regulatory objective remains to be the protection of individual rights respectively basic rights and freedoms of the individual. It does not cover the "structural protection of basic rights," that is, safeguarding the conditions for democratic processes and social inclusion in general by means of guaranteeing individual autonomy in society. In addition, the individual protection of rights and freedoms is not designed to systemically safeguard group-related goals:

The ADM-specific information duties (within the framework of data collection, consent, processing or right to access) includes the duty to provide meaningful information about the logic involved, as well as the relevance of this processing for the data subject. As shown, this duty neither includes the disclosure of deeper decision-making parameters and weightings, the algorithms employed and how they are combined with other algorithms, nor the source code or potential social impacts. The information to be provided – prior to data collection – about the underlying logic of an automated decision is abstract and includes information regarding the basic design, data categories and the decision-making model only. We are therefore not able to thoroughly scrutinize ADM system functionality and, as a rule, it is impossible to detect structural discrimination. In the context of the users' right to information, the data controller is not compelled to justify the decision itself or the system's weightings. Plus, in single-case approaches, there are no automatic conclusions about the possibility of systematic intentional or unintentional discrimination. This would involve having access to a large number of single decisions. The right to obtain human intervention, which is advantageous for the individual, cannot facilitate the verification of a system or process as a whole, or an evaluation of an ADM system that is based on multiple cases. In other words, the transparency needed to assess group- or societal-related risks differs in terms of its depth and extent from the type of transparency specified in Art. 12, 13, 14, 15 and 22 GDPR, the purpose of which is, again, to safeguard individual rights.

At least there are aspects of the GDPR provisions addressing systemic and procedural aspects that can have a positive impact on attaining supra-individual normative goals. In particular, the rules governing the systematic appraisal of privacy-by-design concepts, the duty to compile a list of data-processing activities and the ADM-relevant duty to conduct a data protection impact assessment (not to mention the adoption of internal data protection guidelines), can help establish decision-making and development processes which might also be able to address group risks internally. As we have shown, internal examinations focus on the risks to individual rights. Because system-wide decisions have the potential to violate the rights of a large number of individuals, the data controller is naturally motivated to maintain an overview of potential supra-individual risks throughout the operation. Given his deep insight into the system, the data controller can also discover errors that can affect entire user groups. A systemic perspective, though it remains an internal matter, must therefore target systematic risks, too.

This circumstance makes it necessary to take a short look at the role of data protection authorities: Thanks to its wide-ranging information and access rights, the GDPR provides a governance framework that can clearly improve the quality and motivation of the risk-focused development and implementation of ADM systems among data controllers. The impact of these indirect effects depends on the monitoring and supervisory pressure exerted by the data protection authorities. Like other authorities, supervisory authorities also focus on safeguarding individual rights and freedoms. Given the potential in ADM systems for systematic discrimination with the capacity to affect a large number of individuals and violate their rights, authorities can thus keep an eye out for threats to group interests in general during their data protection audits. Merely, the perspective of data protection authorities does not include supervision in terms of targeting societal goals.

However, the significantly extended tasks and powers of data protection authorities may come in useful here. As specified in Art. 57 (1) b) GDPR, supervisory authorities, in addition to monitoring and enforcing data protection regulations, are also supposed to "promote public awareness and understanding of the risks, rules, safeguards and rights in relation to processing". For authorities this task may, at least in theory, be interpreted as maintaining a

watchful eye on socio-technical developments in general and their potential risks to society (in addition to the risks posed to individual rights) and as publishing (or making public) the knowledge the authorities gain in this area. All of this, of course, must be conducted within the framework of their activities. By taking such a “meta-perspective”, authorities may reach the conclusion after conducting many individual audits that the diversity of ADM systems in various economic sectors is limited and that this endangers additional threats to social inclusion. Here, the authorities, in view of the large number of data processing entities, also depend on information supplied by data subjects. The latter’s right to lodge a complaint as specified in Art. 77 (1) GDPR thus acquires a relevance that extends beyond the legal protection of individual interests.

However, GDPR’s general focus on the single user has consequences in terms of safeguarding the normative goals and protective measures described above. It makes all direct forms of verifiability of the assumptions and assessments manifested in the ADM process logic with regard to their risks for supra-individual goals on grounds of information obligations, users‘ access rights and of system-related assessments difficult. This also applies to the inability to access the databases and decisions for more than individual cases, so that instances of structural discrimination or monoculture trends in the use of ADM systems cannot be systematically detected. As a result, with respect to the aforementioned individual-related risks, we see limited possibilities for a comprehensive external evaluation and for the verification and rectification of complete ADM systems. Given their public-awareness raising activities, data protection authorities may in future take a broader perspective on possible risks to supra-individual goals.

But there remain relevant areas that have yet to be addressed by the law. Because of the GDPR’s focus on the individual and individual rights, legal protections of supra-individual matters (i.e., societal issues) appear to be – at best – of secondary importance. Issues ranging from origin-based discrimination to the reinforcement of gender stereotypes and cases of systematic racism are not addressed by the GDPR and BDSG if an individual is not involved. Neither the GDPR nor the new BDSG facilitate an insight by neutral third parties into the internal ADM logic, its models’ appropriateness and its processes that would allow for the detection of this type of structural discrimination. Any completely neutral verification of group-oriented and societal goals is currently impossible. Moreover, the legal framework in the area of data protection does not provide a systematic overview of the status of ADM system diversity. A difficult problem for the current set of tools – because it calls for a macro-perspective that goes well beyond single cases and across systems – is the ability to assess the social impact of ADM systems: Individual rights protection and respective supervisory tools can neither detect nor mitigate most issues of systematic unfair treatment and discrimination of specific groups of people, which result from a large number of automated decisions (and not from a single case). The same is true for the structural weakening of individuals’ social inclusion, when a large number of the same or similar ADM systems single out the same individuals on the basis of a certain set of data. The idea of a data controller demonstrating “accountability” for an ADM system that would include the system’s societal impact is, as a regulatory approach, unknown to the current data protection approach that is aiming at safeguarding individual rights.

6 How the GDPR could bring benefits: Data protection approaches and instruments for the remaining risk potentials

Data protection law takes its bearings from individual rights and freedoms. Therefore, it is not well-suited to deal with every aspect of normative goals, especially in terms of group-oriented or societal interests. This section therefore focuses on regulatory approaches that may complement the aforementioned data protection measures and can thus better safeguard the relevant interests. The focus stays on tools that aim at optimizing the regulation's impact, at comprehensive ADM system transparency and verifiability, and at safeguarding diversity and rectifiability. Here, a distinction needs to be made between the organizational possibilities within the current GDPR framework on the one hand (chapter 6) and alternative forms of regulation that are not yet covered by data protection law (chapter 7).

Regulatory tools within the GDPR framework mostly derive from what is known as "soft law" and can prove effective in terms of prevention specifically. In co-regulatory settings for instance, businesses might give themselves codes of conduct that specify the contours of rather undetermined legal provisions while accounting for societal interests and various kinds of evaluation. Supervisory authorities' expanded influence is in part related to their capacity to observe ADM systems' impact on societal goals, even if this does not involve any direct legal consequences for the provider. Another possibility consists in the use of GDPR opening clauses in German law.

6.1 Co-regulation: Certified codes of conduct as a support for commercial initiatives

Codes of conduct in line with Art. 40 GDPR could enable industry associations to also provide guidance to data controllers regarding issues of supra-individual freedoms and rights. Art. 40 and Art. 41 GDPR form a co-regulatory structure here: Through these codes of conduct, or rules voluntarily adopted by companies and groups of companies, the industry helps define partially undetermined regulatory provisions – given that they have been approved by the competent authorities. When it comes to regulating complex issues, involving industry in forms of self-regulation can be helpful. Thus, part of the regulatory process is in the hands of those who are subject to regulation which, as a rule, businesses in the industry. This has certain advantages, also with regard to ADM systems: Businesses' know-how is incorporated into the process, which makes quicker and more flexible responses to changes in the field possible. Furthermore – and of particular relevance –, self-regulation can attain goals that are safeguarded by very different national laws. Plus, self-regulation does not have to consider national or EU borders. However, self-regulation often requires a governmental framework to ensure effective impact within an industry. This is often referred to as co-regulation or "regulated self-regulation."

A body accredited (in line with Art. 41 GDPR) by the competent data protection authority then monitors compliance with the code. While codes of conduct have been used in data protection law for quite some time, the GDPR provides for a more supportive framework as well as proper procedures, leading to more genuine incentives to create such codes, e.g. when fines are imposed, businesses that have adopted a code may be treated better.

Having a code of conduct that applies to ADM systems across various sectors or within segment-specific areas would be a way of eradicating some of the unclear issues that have been discussed in this analysis (e.g., with regard to decision-making architectures as defined by the GDPR or the concrete implementation and realization of the rights of data subjects). Moreover, the code could address risks beyond those to individual rights. The scope of codes of conduct limited to specify GDPR provisions and could, if necessary, also contain voluntary commitments which involve external evaluation where societal interests are subject to risks. Here, data protection could be the point of departure in developing further initiatives.

Prior to established industry codes of conduct, the GDPR sees the data protection authorities as relevant actors: Indeed, as stipulated by Art. 40 (1) GDPR, they are expected to promote and support the compilation of codes of conduct in co-operation with the EU and the Member States.

6.2 GDPR: Expanding data protection authorities‘ regulatory options

The legal analysss of the GDPR and BDSG provisions have shown that there are possible starting points for legal innovation within the GDPR which can counter the risks associated with ADM systems. But such innovations which build on norms granting power to authorities, imply the need for action by the appropriate authorities and/or stakeholders. Because of the data protection authorities' ability to inspect ADM systems and because of their legal and technical know-how, data protection authorities are key actors in public debates about ADM systems. And with a view on their broad discretionary mandate, these authorities may hold one of the keys to effective regulatory options.

As described in chapter 5, authorities adopting an open approach to their awareness-raising powers, they can pro-actively include societal interests in their observations and assessments without generating negative legal consequences for individual data controllers. Because their supervisory powers are limited to violations of GDPR provisions, the scope of actions they can take against data controllers on such supra-individual grounds is limited. Nonetheless, they can focus on societal targets specifically and can make the public aware of any negative trends they have identified (e.g., a growing monopolization of ADM systems or algorithms in certain areas of decision-making). In addition, as the only institutions with access to a critical mass of individual complaints cases and proceedings, these authorities can detect any systematic weakening of a specific group's rights through a large number of individual and similar ADM processes. Data protection authorities are thus well-positioned to launch an evidence-based public debate. In areas outside their core supervisory mandate, they would have to comply with their obligation to professional secrecy (Art. 54 (2) GDPR) and the statutory principle of neutrality, which means they could issue general assessments only. They would be prohibited from "naming and shaming" any controllers whose ADM systems do not conflict with legal requirements. The same, in fact, applies to the data controllers. When it comes to conspicuous automated decisions that do not in fact violate the rights and freedoms of an individual data subject, the authority may inform the data controller in formal or informal terms. However, the data protection authority does not have the power to legally request a change in systems or processes in such cases. Nonetheless, data protection authorities have become highly specialized representatives of an (important) interest. For this reason, a forum designed to discuss god practices in ADM governance should also include representatives of other interests.

In addition to supervisory and awareness-raising tasks, the GDPR specifies other organizational powers for data protection authorities: They can require additional duties for providers of ADM systems. As we have seen, automated decision-making is based on the prohibition-exemption principle and brings with it the duty to carry out a data protection impact assessment. However, this applies only to systems in line with Art. 22 (1) GDPR, which has a limited scope of application only (see chapter 3.1). According to Art. 35 (4) GDPR, the supervisory authority can compile and publish a list of the data processing operations which are required to carry out data protection impact assessments. If the supervisory authority were to exercise its decision-making leeway, it could oblige all ADM system data controllers (where necessary) to conduct an impact assessment, including those who would otherwise not fall under Art. 22 (1) GDPR. This way, providers of ADM systems that do not make purely automated decisions or which are not relevant regarding Art. 22 (1) GDPR could also be compelled to conduct data protection impact assessments. It might therefore be possible to establish a verifiable duty to conduct an early internal risk assessment and to ensure compliance with minimum standards. However, determining the extent to which such an expansion would prove proportionate in practice (particularly for SMEs) goes beyond the scope of this report.

6.3 Opening clauses: More restrictive national law requirements

Opening clauses in European law provide additional ways of making certain legal provisions more concrete. The GDPR, too, provides opening clauses for the individual member states. If a state makes use of such a clause, it can deviate from the GDPR provisions or make them more concrete. German legislators have made use of this in the new BDSG, including automated decision-making (see chapter 3.1.2). § 37 BDSG expands the exceptional permissibility of ADM systems with regard to the service provision of insurance contracts. However, certain opening clauses can also be used by EU states for more restrictive purposes:

- By drawing on the opening clause in Art. 9 (2) a) GDPR, member states may introduce legal provisions which would generally make it impossible to give consent to the processing of special categories of personal data in ADM systems in the sense specified under Art. 22 (1) GDPR. This means that in the relevant countries, it would be illegal to give one's consent to the processing of data such as race, genetic or health data and political views in an automated decision-making context. However, this would not exclude the possibility of discrimination via other correlating data subject characteristics.
- It would also be possible on the basis of the opening clause in Art. 37 (4) GDPR to insist on the legal duty to appoint a data protection officer whenever any form of ADM system comes into use, regardless to the requirements specified in Art. 37 (1) GDPR.
- If the broader perspective for the supervisory authorities described above unexpectedly necessitates a legal mandate, an appropriate national regulation could draw upon Art. 58 (6) GDPR to establish such a provision on a national basis.

Thus, opening clauses in the GDPR might entail opportunities for regulating (within narrow limits) ADM more restrictively than foreseen by the GDPR provisions. But the fundamental criticism of the (very large number) of opening clauses could also be applied to the examples cited above: Every national deviation from the provisions of the EU regulation leads to a fragmentation of Europe's data protection framework and thus undermines the purpose of the regulation, which is to harmonize data protection throughout Europe by creating one single legal framework.

This begs the question as to why individual data protection law is invoked as the legal regime that is increasingly concerned with safeguarding the interests of the general public. As discussed, "burdening" data protection frameworks with supra-individual aims often leads to conflicts with traditional data protection doctrines in terms of their protective purposes. In terms of safeguarding normative supra-individual and societal goals, it might be better to use regulatory measures outside of data protection, thus avoid burdening individual-oriented data protection tools with supra-individual protective targets. We need a debate over the proper area for ADM-related regulatory measures that do not focus on data protection of the individual.

The provisions in the GDPR which deal with the relationship between machine-based and human decision-making elements underscore a problem which transcends data protection law. The risks to individuals, groups and society as a whole are associated with the architecture of the decision that creates a relationship between these elements. Whether or not a human being in an ADM process can influence the result – both formally and substantive – depends on several factors determined by the architecture. It is not only the ADM specific law (i.e., data protection law) that significantly regulates the decision-making. Liability law also might be relevant in determining whether it is rational for a human decision-maker to either ignore an ADM system suggestion (as a result of her own assessment) or simply to adopt it. This means that for many of the aforementioned risks it is necessary to analyze all factors that shape a specific decision, and these factors might well be found beyond a single area of law. Modelling decision-making architectures and evaluating different architecture types is a vast field of interdisciplinary research. The findings yielded by this research area should be the subject of intense public debate.

7 Beyond the GDPR: Alternative regulatory tools not covered by data protection law

With regard to the structural limits of data protection law described above, a short discussion of other areas of the law seems pertinent in order to identify potential alternative means of safeguarding the objectives threatened by ADM systems. These alternative tools emphasize different things. They concentrate, for example, on the explainability or verifiability of ADM systems, ensure a diversity across ADM systems or improve their rectifiability. Most notably, they have impact after systems have been implemented already.

7.1 Explainability of automated decisions as a regulatory approach

One of the proposed regulation approaches concerns improvements regarding the legal explainability of ADM systems independent from data protection regulation. These proposals are driven by a variety of objectives that range from data subject rights to answering the "why" question to improving internal protections against faulty – and potentially inefficient – systems to facilitating comprehensive external evaluation.

The issue of ADM systems' explainability creates certain problems which also play a role in data protection law (see chapter 3.4.2), but are of broader relevance. What decisions can be explained and how it can be explained depends on the technical systems in use. Many ADM systems employ stochastic calculations. Using large amounts of data (i.e., training data), these systems establish statistical correlations between (group) features and, for example, the probability of a default risk in a line of credit. From the contoller's point of view, these methods make feasible a variety of assessments regarding the data subject. However, ADM systems do not calculate causal relationships. So-called deterministic systems, in which a specific data input always produces the same output, can (in theory) be programmed to deliver an "explanation" of their decision. However, this kind of mathematical explanation regularly fails to meet human expectations of rationality, causality and consistency that are generally associated with the term. Predictions resulting from calculations using variables in parallel multidimensional processes can be depicted in mathematical terms, but might be exceedingly complex for human interpretation. Thus, depending on the complexity of the system, providing a simple explanation of an individual decision including causal reasoning may not be possible ("curse of dimensionality;" Edwards/Veale 2017:27).

There are also challenges in systems that draw on "artificial intelligence" (AI) in their functionality. The term AI remains subject to definitional debate. Most would agree that the purpose of AI is to increase a system's cognitive efficiency. It also renders these systems more autonomous. In addition, machine learning is increasingly playing a significant role. The systems are programmed in such a way that they can recognize patterns in sets of training data without outside help. In what is known as unsupervised learning there are no ex ante objectives or criteria. The system seeks and identifies patterns entirely on its own. Such continuously learning systems are currently being used primarily for research purposes and (at the moment) to a lesser extent in the products that are being offered to the end user. However, in AI systems, this can lead to a situation in which the developer himself can no longer predict an input's generated result. For systems of this kind, there are three approaches that can help clarify how they function:

- It would be important to determine whether there are any system-specific descriptions (e.g., information about learning-oriented goals fed into the system and information about the training data's composition). This raises the question as to whether the information is ultimately sufficient in order to perform the various risk-control functions (e.g., can a third-party check for system errors and is there a sufficient basis for taking legal measures (see below)). Even simplified models might contribute to better comprehension here.
- ADM systems could be developed in a way to guarantee an "explainability by design," which might involve protocols running in parallel to the ADM process that co-log system changes or run tests informing about the system's status. In informatics, decoupled decision-making and decision-

justification systems are an important area of research. This can be attributed to the growing relevance of decompositional processes (Edwards/Veale 2017) where, within the framework of an ADM system, different modules run in parallel where their various decision-making steps are annotated independently.

- Even in cases of completely explainable automated decision-making, there ought to be a debate about which metrics the testers use or should use. How, for example, can fairness be operationalized not only in mathematical terms, but also against the background of the normative goals? Who decides whether or not the results can be interpreted, and who chooses the interpretative approach? How, within the framework of evaluations, would a legal practice emerge when it comes to harmonizing inspection criteria?

Debates regarding these three aspects have only just begun and will require more transdisciplinary research. Any potential regulatory intervention in the form of a concrete case-related duty to explain automated decisions would involve determining which approach could support the implementation of such explainable systems and, given the alternatives, to what extent this is necessary. Understandably, people want justifiable automated decisions. However, these decisions must also be compared with the rationality of human decisions (cf. Ernst 2017). The implicit expectations regarding the two decision-making models should be the same, yet irrational or hidden decision-making parameters are regular features of human decision-making architectures (e.g., forms of implicit knowledge or an implicit decision-making heuristic). While both human and automated decisions in public administration always require some form of justification, it is inherent to personal autonomy in private contracts that a private individual or body is also allowed to make irrational or biased decisions – as long as the decision does not disregard the anti-discrimination requirements of national law.

7.2 Enhanced transparency and accountability provisions enabling third-party evaluations

The aforementioned obstacles associated with promoting legal objectives through transparency obligations are less relevant where third parties with a profound understanding of the subject and sufficient motivation are the observers. Transparency-related regulatory advantages become more tangible when experts – also outside of authorities – carry out neutral assessments, and especially when they look beyond individual rights and include group-oriented and societal goals. However, for data controllers, whose competitiveness may (also) derive from their use of ADM technologies, confidentiality obligations would seem to preclude any in-depth view of their decision-making systems. If external access to this information is to be regulated by law, these conflicting rights would have to be taken into account, perhaps in strict court-like in-camera proceedings in which the confidentiality of the independent experts is absolutely essential. At any rate, the public documentation of internal processes, decisions and software code, which has occasionally been suggested in policy debates, seems rather disproportionate.

In case of in-camera proceedings, points of reference for disclosure or access permissions would be the various constituent parts of the ADM system such as data sources, data structure and ADM source code, but might extend to the compliance level (e.g., access to (internal) documents compiled as a result of GDPR provisions, including data protection impact assessments, directory lists, internal guidelines or organizational and procedural privacy by design measures). The advantage of in-camera proceedings could be the explicit duty on the part of the experts to consider societal goals in their ADM assessment. The disadvantage would be that the proceedings are not open to the public and that a societal debate, especially with the participation of civil society actors, would be based indirectly on the findings. Moreover, with any kind of institutionalized appraisal by third parties, there is a risk that it may lead to the establishment of supervisory-like procedures in parallel to those of the data protection authorities, where the two will not be sufficiently distinct from one another.

7.3 Options for external evaluation without access to the system

Another approach is seen in current research on model-agnostic evaluations of ADM systems (or “black box testing”). Such deliberations start with the fact that an evaluator of a system either does not have access to the internal workings of the system, or that such access provides no useful insights into possible discriminations by the system. As in the case of reverse engineering that is used to understand software, methods of so-called “pedagogical rule extraction” try to obtain evaluation insights from a systematic interrogation of the systems by using test data. Repeated and coordinated querying enables the interrogation process to come to certain conclusions about the predominant models and weightings incorporated in the ADM system. The evaluator normally queries the ADM system in automated form via an application programming interface (API). In order to facilitate the widespread use of such processes, ADM providers would have to be obliged to procure and install respective APIs. The advantage of pedagogical rule extraction is that an evaluation does not require access to the software code – the very core of the provider’s confidentiality interests. A disadvantage might be that such systems are not easily scalable. In this approach, there is also the question of who pays for what and who, in fact, will take over this cumbersome work.

7.4 Options for application of consumer protection and competition law for improved rectifiability

Competition law is considered to be the faster and more flexible brother of administrative law. Because of the large number of “supervisors” (competitors, that is) and organizations with the authority to take legal action (e.g., consumer protection offices and associations fighting unfair commercial practices), violations of competition law are quickly detected and lead to cease and desist letters. In the case of consumer protection violations, it does not take long to obtain an interim injunction by a court, too. But to what extent competition law and consumer protection law can be used as a regulatory measure for ADM systems would require a far more encompassing exploration, which goes beyond the scope of this report. However, in competition law, the violation of a data protection provision can easily be deemed an unfair commercial practice, leading to a more effective enforcement of existing laws compared to data protection frameworks only, making it more likely to attain the aforementioned goals. The same applies to cease-and-desist orders, in which organizations entitled to sue for an injunction issue a warning to the data controller on grounds of assumed violations of data protection regulations (§§ 3, 2 (1) no. 11 Injunction Act, Unterlassungsklagegesetz (UKlaG)). The application of competition law would remain limited, however, as a result of the incorporation of data protection regulations, which focus primarily on individual rights and freedoms. Furthermore, there would be the de facto problem of whether the plaintiffs who, as a rule, can access only individual cases, can prove that a discriminating ADM system violates data protection legislation on a structural level. Organizations able to take legal action might be able to obtain a better overview in the case of class-action lawsuits where they represent a large number of plaintiffs. Such class-action suits are often discussed in consumer law, though at the moment German competition law does not provide for this instrument. In the case of a civil law complaint, the civil court would have to verify the alleged data protection violation and usually would summon expert advice, at least in the case of complex ADM systems. This means that the form and extent of the supervision in the case of competition law pleas is comparable with those under data protection law, though civil cases may be faster. Here again specific societal risks would not be of legal interest. However, parallel supervision by the courts (on the basis of competition law) might also weaken the GDPR supervision by data protection authorities, since there is no guarantee that the interpretational leeway would be applied in the same way in the both areas.

7.5 Options for application or adoption of regulatory tools from competition law and media law to ensure diversity

Possible consolidation tendencies with regard to the use of specific ADM systems or their underlying algorithmic frameworks pose challenges to pluralism in society. Where fewer ADM providers and frameworks are present, the applicability of antitrust regulations aiming at slowing the process of monopolization comes into question. For antitrust regulations to apply it would need individual providers that have a dominant share of the relevant markets. However, the definition of what constitutes the market is not a trivial issue in the case of ADM systems, since licensable systems often perform an abstract function and might be tailored for and implemented in completely different areas of use. Other systems are exclusively developed within a company for a specific purpose. In such cases, it is impossible to acquire a dominant market share. But even for in-house solutions, developers sometimes make use of third parties' algorithmic frameworks. It might be possible to observe the emergence of oligopolies of frequently used frameworks or software packages which might be relevant in the context of antitrust legislation. However, this would require a market overview which, as noted above, is difficult to establish. In order to achieve an overview of the systems and frameworks currently in use, one would have to rely on information provided by industry sources. But even if one were to introduce a regulatory duty of disclosure here, the application of antitrust law would be restricted to identifying market shares and then to issuing economic counter-measures such as an obligation to provide access or – in a worst-case scenario – unbundling obligations. Thus antitrust law would, in theory, provide a rather weak point of departure for attempts to increase the diversity of ADM systems in the market.

From a regulatory point of view, algorithm-based creation of public spheres as well as the algorithmic management of (public) information flows and personalized services in individual opinion-formation have an impact on societal goals such as diversity and the quest for power. Media policymakers have discussed the need to improve the current limitations when it comes to governing media platforms and information intermediaries. Here, a prohibition of discrimination has also been analyzed (Dankert/Schulz 2016). In terms of the identified risks associated with ADM systems, providers or intermediaries of platforms that employ ADM processes in media content in the purported service of individual interests are of particular. Extending the application of media law requirements to these providers regarding diversity obligations or discrimination bans could prove relevant for automated decision-making designs. This report cannot provide detailed answers to these questions, but it seems likely that even in such cases, diversity requirements and the duty to enforce non-discrimination will refer to the output of ADM-based services rather than how these results are produced. Applying media law thus might help enhance pluralism and non-discrimination in public communication. However, it wouldn't strengthen system-related transparency.

8 Conclusions

Automated decisions bear considerable potential but also pose risks for normative goals regarding the individual such as autonomy, personality rights and fair treatment in terms of non-discrimination. They also pose risks for group-oriented goals of non-discrimination and societal goals like pluralism and social inclusion. The three basic measures for dealing with these risks are transparency, verifiability and rectifiability of automated decisions and the ADM systems on which they are based. When it comes to ADM processes, the data protection framework that is applicable from May 2018 on is limited in its capacity to safeguard these goals. The GDPR's narrow definition of automated decisions and the rather far-reaching exceptions will help create an environment in which interaction with ADM systems will remain an everyday occurrence.

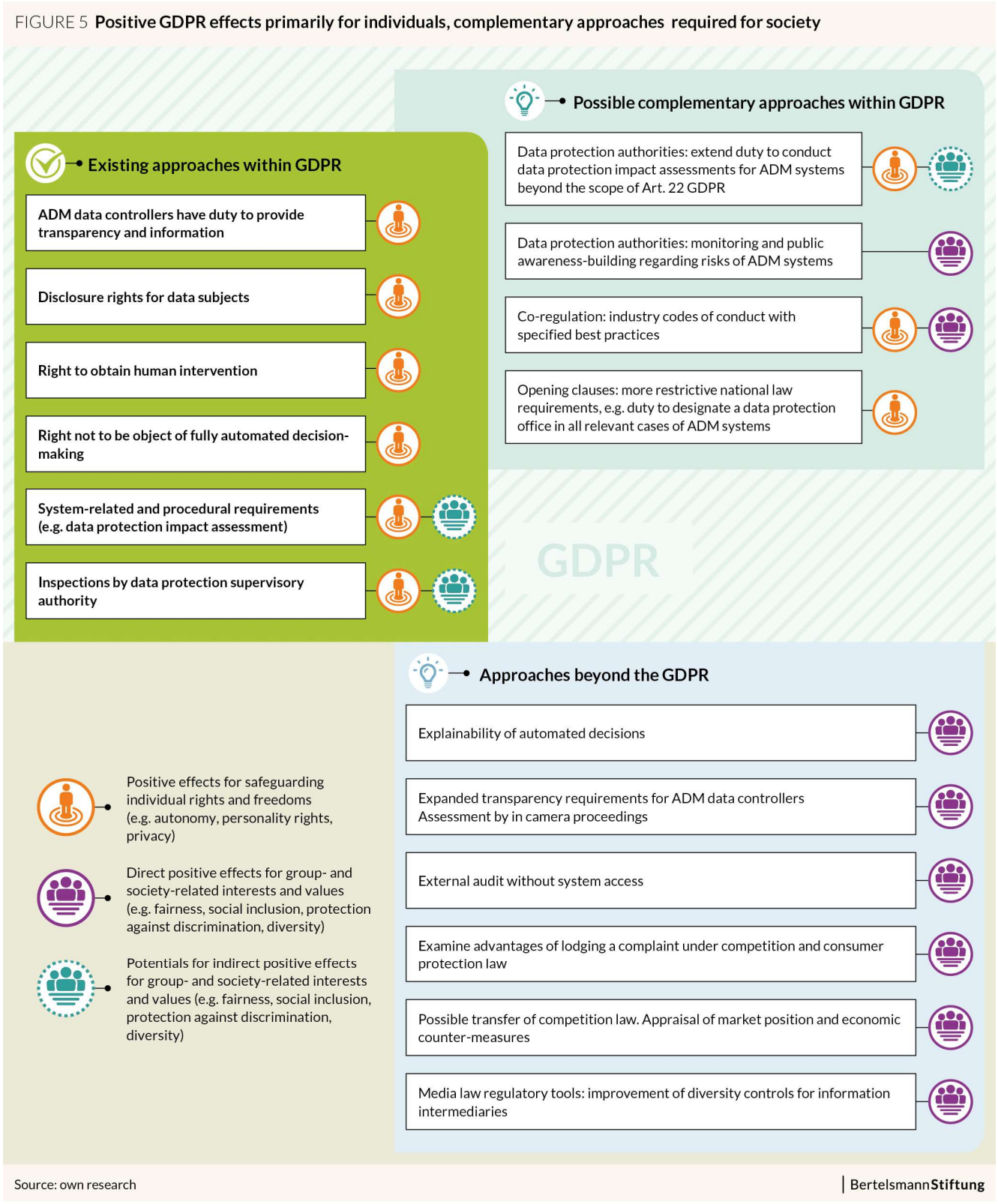
Given the regulatory character of data protection with its focus on the rights and freedoms of individuals, the GDPR offers regulatory tools for single-case transparency, comprehensibility and the rectifiability of an automated decision. In particular, when compared to completely automated processes, the right to obtain human intervention – i.e. a person who can scrutinize the decision, evaluate the issue in conjunction with subsequent statements, and then make a new decision – is able to protect the personality rights of the data subject and her desire to be treated fairly. The data protection framework also provides a duty to provide information about the logic involved, but the scope and depth of the respective information is restricted to a degree of difficulty that the average user can understand. The GDPR thus creates a legal framework which can protect individual interests only, making individual decisions verifiable and rectifiable. The transparency and user rights specified in the GDPR do not provide for deeper insight into ADM systems' functionality that would point to risks of systemic discrimination and negative sociotechnical effects.

Furthermore, the GDPR obliges the data controller with regard to ADM system design and implementation. In addition to individual goals, data controllers might consider – rather incidentally and to a limited extent – group-related issues such as non-discrimination, e.g., during data protection impact assessments or by applying privacy by design principles during the design phase. Ultimately, the regulation manages to strengthen awareness among ADM system providers regarding data subjects' rights and freedoms. However, the legal framework is of limited use when it comes to structurally safeguarding group- and societal-related goals.

The GDPR grants data protection authorities useful powers and instruments in terms of extended information rights as well as access and inspection powers. Supervisory authorities are thus theoretically positioned to strengthen the providers' mentioned duties. In the event of a data protection audit, authorities will have full access and the capacity to verify the processes and structures used by a provider including any impact assessments. While authorities are obliged to primarily focus on the protection of individual rights in such cases, they might expand the reach of their monitoring function by carrying out awareness-raising activities across single audits in parallel. Authorities could also point to undesirable structural developments beyond audits of individual providers. In public debate, data protection authorities are relevant actors when it comes to addressing the societal risks posed by ADM systems. The extended catalogue of far-reaching powers contains several potential paths to pursue here, though this would involve a more pro-active role on the part of authorities. However, the GDPR does not specify a right to ADM systems access and evaluation by independent third parties such as scholars or technical experts. As a result, relevant measures like system-related transparency, ADM system verifiability and safeguarding diversity are absent in data protection frameworks. We need to pursue regulatory debates that go beyond data protection law.

Against this background, this report points toward potentially effective ways forward in data protection law and other areas of the law which, if properly implemented, might stimulate debate and better safeguards societal norms and values.

Fig. 5: Positive GDPR effects primarily for individuals, complementary approaches required for society



Literature

Article 29 Working Party (2017). ”Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679“. WP 251, 3rd October 2017. Available at http://ec.europa.eu/newsroom/article29/document.cfm?doc_id=49826 (last visited on 7.3.2018).

Bäcker, Matthias (2017). “Art. 15. Auskunftsrecht der betroffenen Person“. *Datenschutz-Grundverordnung: DSGVO. Kommentar.* 1. ed. Hrsg. Jürgen Kühling and Benedikt Buchner. Munich. 387–397.

Barocas, Solon, and Andrew D. Selbst (2016). “Big Data’s Disparate Impact“. *Californian Law Review* (104). 671–732.

Bertelsmann-Stiftung (ed.) (2017a). *Wenn Maschinen Menschen bewerten – Internationale Fallbeispiele für Prozesse algorithmischer Entscheidungsfindung.* Gütersloh. (Available at https://algorithmenethik.de/2017/05/02/wenn-maschinen-menschen-bewerten/, last visited 8.2.2018.)

Bertelsmann-Stiftung (Hrsg.) (2017b). *Digitale Öffentlichkeit. Wie algorithmische Prozesse den gesellschaftlichen Diskurs beeinflussen.* Gütersloh. (Available at https://algorithmenethik.de/2017/06/19/digitale-oeffentlichkeit-wie-algorithmische-prozesse-den-gesellschaftlichen-diskurs-beeinflussen/, last visited 8.2.2018.)

Bull, Hans Peter (2006). “Zweifelsfragen um die informationelle Selbstbestimmung – Datenschutz als Datenaskese?“ *NJW (Neue Juristische Wochenschrift)* 2006. 1617–1624.

Bygrave, Lee A. (2001). ”Minding the Machine: Article 15 of the EC Data Protection Directive and Automated Profiling“. *Computer Law & Security Report* (17). 17–24.

Edwards, Lilian, and Michael Veale (2017). ”Slave to the Algorithm? Why a ‘Right to an Explanation’ Is Probably Not the Remedy You Are Looking For“. *Duke Law & Technology Review* (16) 18. 18–84.

Ernst, Christian (2017). “Algorithmische Entscheidungsfindung‘ und personenbezogene Daten“. *JZ (Juristische Zeitung)* 21. 1026–1036.

Gola, Peter (ed.) (2017). *Datenschutz-Grundverordnung: DS-GVO. Kommentar.* 1. ed. Munich.

Goodman, Bryce, and Seth Flaxman (2016). ”European Union regulations on algorithmic decision-making and a ‚right to explanation‘“. arXiv:1606.08813 [stat.ML]. https://arxiv.org/pdf/1606.08813.pdf (last visited 7.3.2018).

Howells, Geraint G. (2005). ”The Potential and Limits of Consumer Empowerment by Information“. *Journal of Law and Society* (32). 349–370.

Keats Citron, Danielle, and Frank Pasquale (2014). ”The scored society: Due Process for automated predictions“. *Washington Law Review* (89). 1–33.

Malgieri, Gianclaudio, and Giovanni Comandé (2017). ”Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation“. *International Data Privacy Law*, ipx019. https://doi.org/10.1093/idpl/ipx019 (last visited 7.3.2018).

Martini, Mario, and David Nink (2017). “Wenn Maschinen entscheiden ... – vollautomatisierte Verwaltungsverfahren und der Persönlichkeitsschutz“. *NVwZ (Neue Zeitschrift für Verwaltungsrecht)* (36) 10. 1–14.

Mendoza Isak, and Lee A. Bygrave (2017). ”The Right Not to be Subject to Automated Decisions Based on Profiling“. In: Tatiani Synodinou, Philippe Jougleux, Christiana Markou and Thalia Prastitou (eds.), *EU Internet Law: Regulation and Enforcement.* Cham. 77–98.

Radlanski, Philip (2016). *Das Konzept der Einwilligung in der datenschutzrechtlichen Realität.* Heidelberg.

Schulz, Wolfgang, and Kevin Dankert (2016). *Die Macht der Informationsintermediäre. Erscheinungsformen, Strukturen und Regelungsoptionen.* Bonn.

Selbst, Andrew D., and Julia Powles (2017). "Meaningful information and the right to explanation". *International Data Privacy Law* (7) 4. 233–242.

Wachter, Sandra, Brent Mittelstadt and Luciano Floridi (2017). "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation". *International Data Privacy Law* (7) 2. 76–99.

Weil, David, Archon Fung, Mary Graham and Elena Fagotto (2006). "The Effectiveness of Regulatory Disclosure Policies". *Journal of Policy Analysis and Management* (25) 1. 155–181.

Zarsky, Tal Z. (2017). "Incompatible: The GDPR in the Age of Big Data". *Seton Hall Law Review* (47) 4. 995–1021.

Zweig, Katharina A. (2018). "Wo Maschinen irren können. Fehlerquellen und Verantwortlichkeiten in Prozessen algorithmischer Entscheidungsfindung". Available at https://algorithmenethik.de/2018/02/05/wo-maschinen-irren-koennen-fehlerquellenund-verantwortlichkeiten-in-prozessen-algorithmischer-entscheidungsfindung/ (last visited 8.2.2018)

About the authors

Prof. Dr. Wolfgang Schulz is director of the Hans Bredow Institute for Media Research at the University of Hamburg. In November 2011, he was appointed Professor of Media Law and Public Law at the Faculty of Law of the University of Hamburg. The professorship is funded jointly by the University of Hamburg and the Hans Bredow Institute, and focuses on the research being conducted at the Hans Bredow Institute. He was initially deputy managing director and head of the department of media and telecommunications law. In July 2001, he was elected to the Board of Directors. In February 2012, he was appointed director of the Alexander von Humboldt Institute for Internet and Society in Berlin. He is also a member of the Committee of Experts on Internet Intermediaries (MSI-NET) of the Council of Europe. Wolfgang Schulz studied law and journalism in Hamburg. Since 1997, he has lectured on information and communication at the Faculty of Law of the University of Hamburg. In January 2000, he became a member of the State Legal Examinations Office. In July 2009, he was awarded a habilitation degree by the Faculty of Law of the University of Hamburg, entitling him to teach public law, media law and legal philosophy.

His research focuses on the freedom of communication, problems of legal regulation with regard to media contents, questions of law in new media, above all in digital television, and the legal bases of journalism, but also the jurisprudential bases of freedom of communication and the implications of the changing public sphere on law. In addition, he works on the forms taken by the State’s functions, for instance, in the framework of concepts of “regulated self-regulation” of "informational regulation". Many of his projects are designed internationally comparative.

Stephan Dreyer is Senior Researcher for Media Law and Media Governance at the Hans Bredow Institute for Media Research, Hamburg.

His research focuses on regulatory issues of mediated communication in a datafied society. He looks into challenges that regulation is facing in the light of new technologies, services and changing media use. Currently he is working on legal issues of AI-based communication and automated decision-making systems, (social) bot communication and the limitations of transparency/disclosure as a regulatory resource. He's a legal expert in regulatory questions at the intersection of protection of minors, privacy and data protection protection and also conducts legal and comparative analyses of systems in the area of media-related governance. He has been a research associate at the Hans Bredow Institute since 2002. Stephan Dreyer is a law graduate and studied law with special emphasis on information and communication at the University of Hamburg.

Stephan Dreyer is legal spokesman of the Complaints Committee and the Committee of Experts of Freiwillige Selbstkontrolle Multimedia-Diensteanbieter e.V. (FSM) and youth protection expert at the USK.online. He is a founding member of the "Center for Social Responsibility in the Digital Age" (SRDA).

Adresse | Kontakt

Bertelsmann Stiftung
Carl-Bertelsmann-Straße 256
33311 Gütersloh
Telefon +49 5241 81-0

Ralph Müller-Eiselt
Director
Telefon +49 5241 81-81456
ralph.mueller-eiselt@bertelsmann-stiftung.de

www.bertelsmann-stiftung.de

BertelsmannStiftung