De Modernisering van het Nederlands Procesrecht in het licht van Big Data:

Procedurele waarborgen en een goede toegang tot het recht als randvoorwaarden voor een data-gedreven samenleving

Bart van der Sloot & Sascha van Schendel

English summary

Big Data and data-driven applications are increasingly used in both the public and private sector. Banks and insurers use risk profiles, whereby risk groups are identified on the basis of statistical correlations, internet companies use profiles for personalizing advertisements, search results and news items, and Fitbits, total genome analysis and personalized medicine are an integral part of the health domain. Governments around the world are also investing heavily in data-driven processes, such as intelligence and security services that collect and analyze large quantities of metadata, Tax Authorities that are transforming in data-driven organizations and the law enforcement agencies that experiment with predictive policing. In addition to the use of Big Data in the private and public sector, there are numerous private-public partnerships in which Big Data plays a major role, such as the many smart cities and living labs, in which data collection, data analysis and predictive interventions are used to make the public space cleaner, safer and more efficient.

Big Data is used in more and more sectors of society, a trend that is likely to increase in the coming years. To ensure that Big Data is properly embedded in the public sector, it is important that, in addition to focusing on potential benefits, such as a more efficient and effective government, the various risks and bottlenecks associated with Big Data are addressed and mitigated.

Up to now, particular attention has been paid to the protection of the substantive rights of citizens and the emphasis has generally been on principles of material justice. For example, the General Data Protection Regulation grants new rights such as the right to be forgotten and the right to data portability. In addition, both literature and case law is concerned with the potential negative consequences of Big Data, such as the chilling effect, the limitation of individual freedom through nudging, the Matthew effect, the transparency paradox, the filter bubble and the dangers associated with discrimination, stigmatization social stratification. Issues related to material justice are essential for the integration of Big Data in the public sector, as is granting strong substantive rights for citizens

Equally important, however, is that sufficient attention is paid to issues related to access to justice and principles of procedural fairness. Citizens who have rights but are unable to successfully enforce them remain empty-handed and a legal system that addresses incidental Big Data harms only on an individual level does not tackle the underlying causes, so that structural problems may persist. Issues related to procedural fairness and access to justice vis-à-vis Big Data projects have received little attention so far. This is remarkable, not only because this has left a number of complicated dilemmas underexposed, but also because Big Data raises a number of new legal challenges that are precisely related to these points. This report addresses those questions, analyzes the essential characteristics of procedural law and indicates where improvements are possible.

Both national and supranational procedural law is currently characterized by a strong focus on protecting the interests of natural persons in the specific circumstances of a concrete case; in particular, it offers legal protection by granting subjective rights to individuals with a direct and personal interest. This emphasis works well for many traditional legal disputes: a request for compensation after a defamatory publication, a building permit denied by the local municipality or a restriction of the right to privacy by the government, where someone's

phone is tapped for a certain period or someone's house is entered by the police. In each of these cases, the potential infringement or effect is limited to a specific person or a small group, the possible violation can be delineated in time and space, and the interests at stake are clear and directly linked to natural persons.

This is different for modern human rights issues that revolve around large data-driven processes. It is quasi impossible to delineate Big Data projects in terms of time, space and person, as they are a structural and integral aspect of the actions of government services, companies and citizens. The cameras on the corner of almost every street in large cities, for example, do not have an effect that is specifically related to a particular individual; rather, they permanently film everyone in the city; an intelligence agency that collects the communication data of an entire neighborhood or city does not affect anyone specifically or individually, but rather everyone equally; law enforcement authorities that use predictive policing based on postal codes to monitor certain neighborhoods more than others do not harm specific individuals, but target certain communities and may therewith reinforce inequality in society.

The larger the data-driven processes and the more general the data analysis, the more difficult it will be for an individual to substantiate a concrete, direct and personal interest. In essence, Big Data processes do no so much affect individual interests, but affect collective and general interests. Big Data merits reflection on a more general level. Do we want a society in which the public space is constantly monitored and in which authorities can experiment with behavior modification of its citizens? What are the appropriate constitutional safeguards in relation to potential abuse of power by public authorities and how can democratic legitimacy be ensured in public-private partnerships? What is the impact of personalizing insurances and social security rights on solidarity in society?

Besides such matters of general interest, procedural concerns are important in Big Data processes: aspects that relate to how systems and processes are structured, what choices have been made and what consequences those choices have. How are data collected, by whom and where and what influence do such choices have on the possible bias in a dataset with regard to neighborhoods or groups in society? Which standards apply to designing algorithms, the analysis of datasets and the statistical correlations found in terms of quality, accuracy and reliability? And what guarantees are there that the data-driven processes are transparent and auditable and that the deployment based on data-insights is subject to adequate supervision?

Although some legal doctrines exist in both national and supranational law to address these types of questions, such possibilities are limited and do not account for the fundamental changes that are brought about by the data-driven environment. This study looks at which adjustments could ensure a better and more robust embedding of Big Data in the public sector, whereby general and social interests are safeguarded, stakeholders can effectively assert their rights and the principles of procedural justice are respected.

Chapter 2 discusses procedural law in the domains of Dutch civil law, administrative law and criminal law. The conclusion is that, although there are some possibilities in every domain of law to raise points of general interest, they seem insufficient to adequately address the issues that Big Data brings about. That is why Chapter 3 examines which procedural alternatives exist in other jurisdictions that could serve as a source of inspiration. Case law of the European Court of Human Rights of the Council of Europe, legislation of the European Union and the case law of the European Court of Justice were examined, as well as the

legislation and legal practice in Belgium, Germany, France and the United Kingdom.

To get a better picture of the obstacles and bottlenecks with regard to the current legal system and the potential alternatives, interviews were held with a number of key players in the field of Big Data and legal protection: Amnesty International Netherlands (Doutje Lettinga & Nine de Vries), Data Protection Authority (Aleid Wolfsen), Tax Authority (Raymond Kok), Bits of Freedom (David Korteweg), Boekx lawyers (Otto Volgant & Charlotte Hangx), Bureau Brandeis (Christiaan Alberdingk Thijm), Netherlands Institute for Human Rights (Jan-Peter Loof & Juliette Bonneur), Data Trade Union (Reinier Tromp), Supreme Court (Ybo Buruma), State Attorney Pels Rijcken (Cécile Bitter), National Ombudsman (Reinier van Zutphen & Martin Blaakman), Privacy First (Vincent Böhre), Public Interest Litigation Project (Jelle Klaas), Radboud University (Roel Schutgens & Joost Sillen) and writer/publicist (Maxim Februari). The appendix under 6.1 of this report contains the reports of these interviews.

Two workshops were organized, one in The Hague and one in Brussels, in which a number of prominent specialists gave a presentation and a group of Dutch and European policymakers were subsequently invited to discuss the implications of the findings of this report for national and supranational law and the advantages and disadvantages of the regulatory opinions proposed in this study. During the workshop in The Hague, Marlies van Eck (Leiden University), Vincent Böhre (Privacy First), Doutje Lettinga (Amnesty International Netherlands), Otto Volgant (Boekx lawyers) and Phon van den Biesen (Van den Biesen Kloostra Lawyers) gave a presentation. During the workshop in Brussels, Ianika Tzankova (Tilburg University), Wojciech Wiewiórowski (European Data Protection Supervisor), Marc Rotenberg (EPIC) and Max Schrems (NOYB - European Center for Digital Rights) were invited to kick off the discussion. The appendix under 6.2 of this report contains the reports of the workshops.

Thirteen regulatory options have been formulated in Chapter 4 of this study to ameliorate Dutch procedural law and to make it fit for the data-driven society. These regulatory options can be divided into three clusters. Firstly, the Big Data process itself must be properly regulated in order to make sure that procedural safeguards are guaranteed with respect to both the collection, analysis and use of data. Secondly, access to justice vis-a-vis data-driven projects must be guaranteed and the possibilities for litigants to defend the collective and general interests involved though administrative, civil and criminal law can be strengthened. Thirdly and finally, it is important that there is a good system of checks and balances within the *trias politica*, so that the three powers of government can adequately control each other in the Big Data context.

Although these regulatory options, which are summarized below, are aimed at the Dutch legal context specifically, many of those will have a broader significance and may be of interest for experts working in other jurisdictions as well.

Collection	→	Analysis	→	Use

Big Data processes can be characterized on the basis of three phases: collecting data, analyzing the collected data and using the results of those analyses. Big Data technologies can work with extremely large data sets using smart computers and self-learning algorithms. The analysis of the data is usually aimed at finding general characteristics and correlations, is usually based on statistics and the data are analyzed on a high aggregate level. The correlations obtained with Big Data analysis can be used for all types of predictive, preventive and proactive policy choices.

To be able to test Big Data processes and the consequences for the judiciary system, one can look at the regulation of each of the three phases of Big Data. The earlier in the process potential problems and obstacles are addressed, the sooner negative consequences are tackled and the fewer legal cases about any irregularities will be brought at the end. For example, if the way in which the data is collected and analyzed is biased, this may have serious consequences for the deployment of governmental power when based on the insights from data analytics. However, it is by no means always clear in the third phase of Big Data how data were collected and analyzed and it is often difficult for potential victims of biases to show where potential biased choices were made in the process, not least because both the phase in which data are collected and in which they are analyzed are generally characterized by their lack of transparency.

Regulatory Option I: Establish rules through which the gathering of data, other than personal data, are regulated

Part of the data collection process is currently regulated, but another part is not. The part that is regulated, in particular by the General Data Protection Regulation (GDPR), concerns the data that can be qualified as personal data, i.e. data that can identify someone. However, Big Data processes do not revolve around personal data, they usually involve analysing large amounts of aggregated data that are not related to identifiable individuals. Whether or not a given data point could be considered personal data when it was collected does not really matter; in the phase in which the data are analysed, all data are aggregated and are no longer personal data. Due to the current regulatory approach, part of the Big Data process remains unregulated in the collection phase, namely when non-personal data are gathered. This is important because the GDPR not only offers protection to individuals: it lays down general duties of care for data controllers, specifying for example that data must be relevant for the purpose for which they are collected, that the data must be correct and up-to-date and that data processing must be transparent and subject to supervision. An option could be for the regulator to lay down a number of minimum rules for the collection of data, not being personal data, inspired by the GDPR.

Regulatory option II: Regulate the analysis of data

The phase in which data are analysed is now virtually unregulated, both because the GDPR sets hardly any rules on the analysis of data and because, as mentioned above, data analytics generally does not involve personal data, but large data sets of aggregated data. This means that there are hardly any legal standards for, or legal supervision of, how profiles are made, conclusions about the patterns and statistical correlations are drawn, and how data is analysed. This may be problematic, because if there is an error or a bias in a data set, an algorithm or a profile, this means that government action based on that profile (in the third phase of Big Data processes) will also be biased.

By regulating the second phase, Big Data processes are made more transparent and made subject to quality requirements, thereby preventing or limiting any problems in the third phase. Increasing transparency and monitoring this phase can also lead to a reduction in the number of litigation cases against Big Data projects; in interviews conducted for this study, litigants have indicated that due to the lack of transparency with regard to Big Data projects, they often first need to issue information requests before they can decide whether there is a legal problem that would merit a substantive legal procedure. A law or code of conduct regulating the analysis of data in Big Data processes, which is typically based on statistical principles and correlations, is needed. Inspiration could be drawn from the existing standards and guidelines for statistical

Regulatory option III: Set a sunset clause for Big Data projects

Of the three phases of Big Data processes, the third and final phase, in which the insights of data analytics are used for policy making, is currently regulated best. When data analyses are used, they have an effect on society, groups or citizens. Many existing doctrines can be invoked when such an effect materialises, such as the right to a fair trial, the prohibition of discrimination and the freedom of expression and movement. In general, these doctrines are sufficiently equipped to tackle potential problems in the Big Data era, although the question is to what extent access to the justice and the possibilities to enforce these rights are too (this question will be addressed in the following regulatory options).

What would improve the Big Data process with regard to the third phase is a better assessment of the effectiveness of data-driven processes. Studies show that some of the Big Data initiatives are hardly more effective, if at all, than processes that require no or much less data processing. In such cases, terminating Big Data processes would contribute both to an efficient and effective government and to the protection of the substantive rights of citizens and general and social interests. To this end, a sunset clause could be introduced for Big Data projects as a standard: the project then gets a fixed number of years to prove itself in terms of effectiveness.

Civil law	Administrative law	Criminal law

A second way to look at whether there are sufficient procedural safeguards in the context of Big Data is to zoom in on possibilities in civil law, administrative law and criminal law.

Because Big Data processes usually do not involve a violation of the specific interests of natural persons in an individual case, but concern issues related to social values, primary focus should be on strengthening the possibilities for taking actions in the collective and general interest. Legal persons, such as foundations and associations that are founded for the protection of human rights and societal values, typically take the initiative in such campaigns.

Both in the literature and in the interviews conducted for this study, civil law is regarded as the area of Dutch law with the best possibilities for taking actions in the collective or general interest. Most actions against data-driven projects of the government, such as concerning the storage of fingerprints, the exchange of data by intelligence services, data retention and the revision of the law regulating the intelligence agencies were brought under civil law. Only a small number of improvements are possible here. Criminal law, however, currently stipulates hardly any possibilities for addressing these types of interests and within administrative law, the possibilities to raise more abstract issues that are involved with Big Data processes are limited. The latter, in particular, is critical because administrative law takes an increasingly important position in the legal practice of Big Data projects.

Regulatory Option IV: Strengthen the private enforcement system by making it easier to litigate for societal or collective interests

Although within civil law there are currently adequate possibilities for taking actions in the collective or general interest, a number of minor adjustments can ensure that procedural law is brought in line with the transformation to a data-driven society. Examples include the relaxation of the requirement for legal persons to include in their statutes the general interests they wish to defend in a court case and of the requirement of prior consultation with the party against whom an action in the public interest is being initiated, or the establishment of a fixed amount for non-material damage as a result of unlawful Big Data projects, which would help to cover the costs for such actions, or, finally, to expand the opt-out system for collective actions. For the latter, a first step has already been taken in the form of the Act on the settlement of mass damages in collective action that was recently adopted. Nevertheless, even after the adoption of this law, bottlenecks remain, such as with respect to the costs and the possibilities for financing actions in the collective and general interest.

Regulatory option V: Clarify in which cases discrimination in Big Data applications can or should lead to exclusion of evidence or lowering of a sentence

Regulatory option VI: Expand the possibilities for legal entities to participate in the criminal trial

Dutch criminal law currently has virtually no possibility for taking public interest actions and addressing more structural and general problems. That may in time become problematic because Big Data is increasingly being used by law enforcement authorities and this trend is likely to continue in the future. For example, the types of areas and neighbourhoods that are designated as risk areas by predictive policing systems can have an important impact on where additional patrols are deployed and the types of crimes and perpetrators that are registered. There are limited possibilities for defendants to address the question of whether such a system is biased before a criminal court. This bottleneck could be tackled in various ways. First, possible errors and biases in systems could result in the exclusion of evidence or a reduction in sentences. Second, it is conceivable to give legal persons a bigger role in the proceedings. For example, legal persons could be allowed to join the proceedings against defendants so as to contribute with particular knowledge and expertise, and to potentially argue that a Big Data process underlying an individual criminal case is biased or is suffering from problems on a structural level.

Regulatory option VII: Introduce the figure of a special advocate

In addition, a problem associated with Big Data processes is that the process of collecting, analysing and subsequently using data is often not transparent. As far as possible, governments should be transparent. However, in certain contexts, secrecy can be legitimate. If the intelligence service, law enforcement agencies or the tax authorities would make public how they work, which data they use and how data points are weighed, potential criminals could take this into account. In cases where the government legitimately relies on secrecy in the public interest, for example for purposes of national security, a citizen may not have access to information that forms the basis of his criminal trial or an administrative decision concerning him or her. For such situations, a partial solution for the lack of transparency can be found in the introduction of the figure of a special advocate. This lawyer could review the algorithms, data and underlying documents on behalf of the citizen and conduct the defence with respect to these aspects, but at the same time, would be bound to secrecy and would not share this information with the citizen or with others.

## Regulatory option VIII: Open possibilities to appeal against generally binding rules

Dutch administrative law only allows for appeal against decisions by governmental organizations and not against what are called generally binding rules, such as policy decisions that have no direct impact on concrete cases or specific individuals. Many government actions and choices with regard to Big Data processes in the first and second phase, that is where data are collected and analyzed, will not have a direct effect on citizens and can consequently not be challenged. Nevertheless, it may be important to be able to address potential issues in these stages of the process, such as when data is collected disproportionally or analyzed by biased algorithms, because such errors may have important implications for discrimination, stigmatization and social stratification when using the data and outcomes of data analysis. To make such possible, the prohibition to appeal generally binding rules should be exempted, at least for public interest actions in the Big Data context.

Regulatory option IX: Expand preliminary procedures

Regulatory option X: Expand the role of amicus curiae

Regulatory option XI: Create a litigation fund for the Big Data context

Then there is a more practical point. Because litigation vis-à-vis Big Data projects often involves general and social issues, specialized legal entities will play an important role. This research has shown that one of the main obstacles for public interest litigation is related to the costs involved. A number of solutions is conceivable for this practical obstacle. One way in which costs can be reduced and access to justice could be improved is by expanding the possibilities for asking preliminary questions. In this way litigants get a direct ruling from the highest court; through a relatively short procedure, with relatively few costs, an answer can be obtained to fundamental constitutional questions that are spiralled by Big Data processes. In addition, the role of amicus curiae participation could be expanded. By allowing specialized (legal) entities to put forward arguments and documents in proceedings without being a party to the proceedings themselves, their knowledge and expertise can be used in cases while they do not have to bear the costs of the proceedings themselves. Finally, the costs of general interest litigation could be covered through a fund financed by general resources.
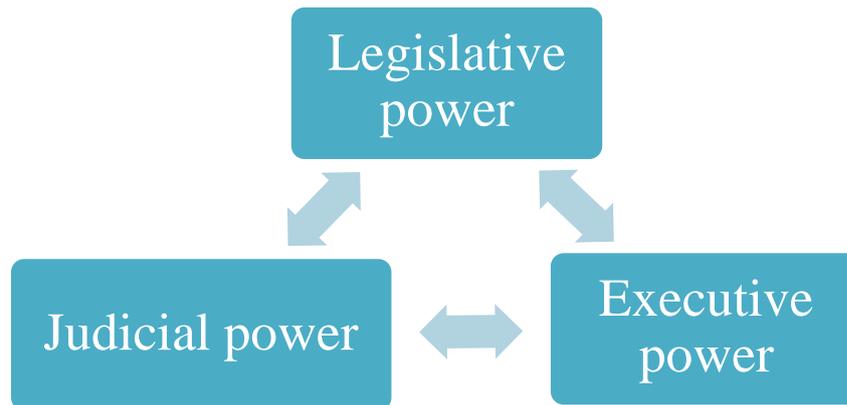
Most of the Big Data processes within the government are deployed by the executive power. Law enforcement agencies, tax authorities, intelligence services and social security agencies, among others, are experimenting with Big Data projects to optimize their processes and work more efficiently and effectively. The first three regulatory options concerned suggestions for legislation and policy to better regulate the Big Data process itself. This was followed by a number of regulatory options to optimize access to justice and to ameliorate procedural law in the data-driven society. For a good system of checks and balances, in which the three powers of government can keep each other in check in the data-driven society, it is not only important that the legislature sets sufficient rules for the executive and that it is sufficiently possible to have the actions of the executive power assessed by the judicial power.

Judicial review of bills and laws and the capacity of supervisory organizations must also be included in this interplay. Two points emerge from this study. First, it is important that the more the government makes use of data-driven processes for the performance of its tasks, the more the capacity for audits and control by supervisory authorities should be expanded. Second, in the Netherlands there is currently almost no room for reviewing legislation *in abstracto*; a system is lacking through which the judiciary can assess whether bills and laws accord to the minimum conditions of legitimacy and legality, while the European Court of Human Rights increasingly demands such from Council of Europe Member States, especially with regard to legislation on large-scale data processing initiatives.

Regulatory option XII: Expand the competencies of supervisory organizations and oversight mechanisms

Supervisory authorities such as the Dutch Data Protection Authority have an important role in handling complaints. Every complaint that a supervisory authority handles satisfactorily will not end up in court. In addition, an organization such as the Data Protection Authority is ideally equipped to investigate and, where necessary, sanction problems that materialize on a more structural or societal level. The Data Protection Authority also has the expertise to ensure compliance with rules under regulatory options I and II. It is therefore important that this authority and other supervisory organizations that have competence vis-à-vis particular aspects of the Big Data process are sufficiently equipped to perform their supervisory duties.

Regulatory option XIII: Create a possibility for *in abstracto* assessments of laws and policies by the judiciary

In court cases involving Big Data processes, one of the most common questions is whether the underlying policy or legislation as such is lawful and legitimate. Legal questions include whether the revision of the law regulating the secret services complies with the human rights framework, whether data retention legislation constitutes a disproportionate violation of the right to privacy and whether the predictive policing initiatives are sufficiently transparent. In the Netherlands, the possibilities for testing legislation as such on its legality and legitimacy are limited, while the ECtHR assumes that Member States to the Council of Europe allow such and otherwise will take matters into its own hands. Not only does the introduction of such a possibility have the advantage that laws and policies are tested directly by national courts, meaning generally shorter procedures. In addition, the number of cases may be reduced by introducing possibilities for *in abstracto* assessments by the judiciary, as not every individual who believes to be negatively affected will submit a complaint individually. A court can assess the law or policy as such in relation to the principles of legality, legitimacy and the rule of law, even before they are applied.

These are the thirteen regulatory options that emerge from this report. They are worked out in more detail in Chapter 4. They cover the regulation of Big Data processes as such, procedural safeguards and access to justice vis-à-vis data-driven processes, and the checks and balances needed in the data-driven society. Each of the regulatory options can be seen as a policy recommendation, but that is not the primary function of identifying these regulatory options.

Not all regulatory options will be feasible in the short term and some regulatory options can be seen as communicating vessels. For example, if there are good possibilities within administrative law to address the various problems that can arise with Big Data processes, less has to be done to strengthen the possibilities within criminal law and civil law.

In that light, it is particularly important that the legislator assesses the pros and cons of all different options, taking account of how the various options could be embedded in the legal system.

Ultimately, the goal should be redesigning the legal system in such a way that it offers citizens and society as a whole sufficient legal protection in the data-driven society.