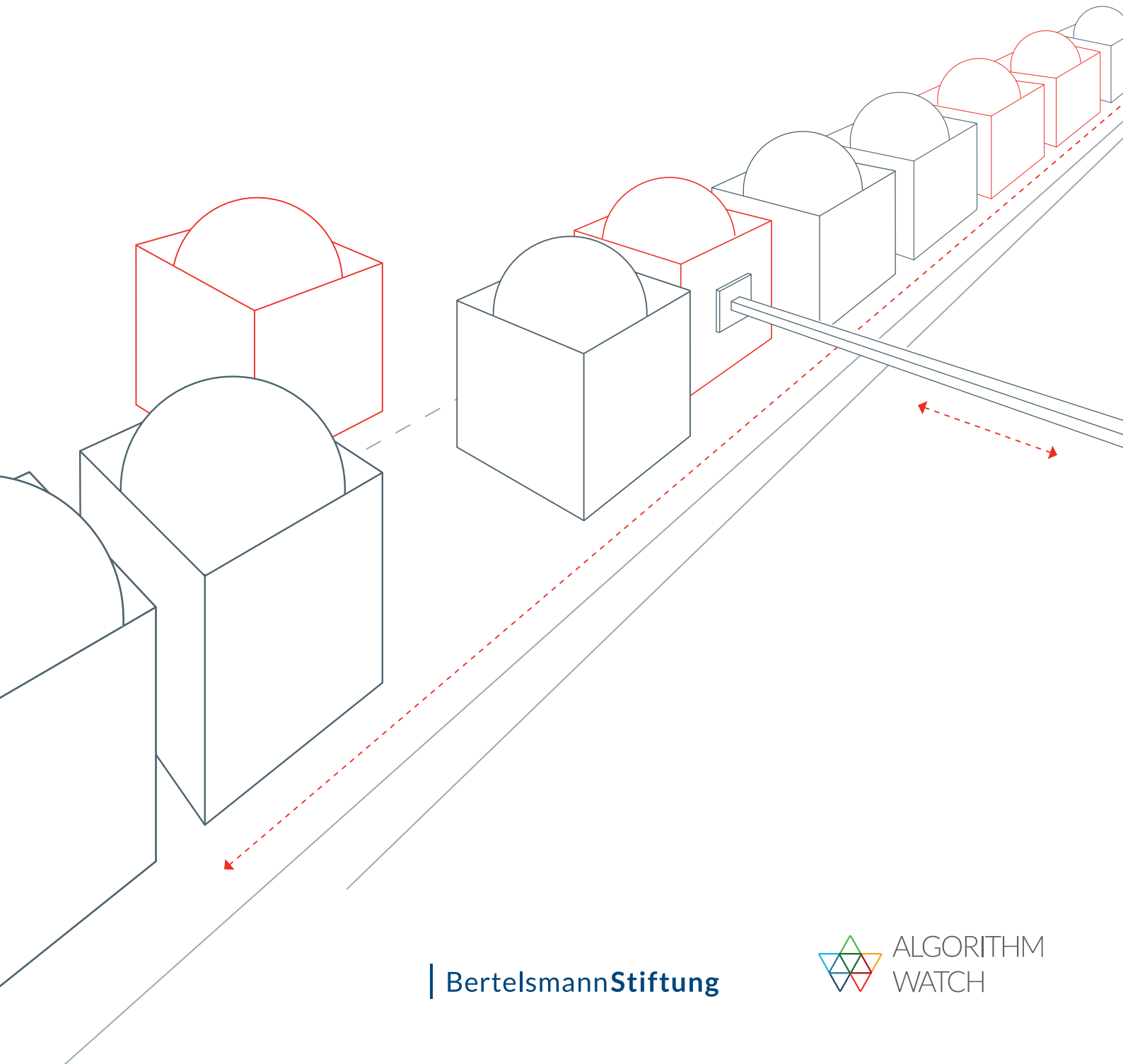# Automating Society

Taking Stock of Automated Decision-Making in the EU

A report by AlgorithmWatch in cooperation with Bertelsmann Stiftung, supported by the Open Society Foundations

Automating Society
Taking Stock of Automated Decision Making in the EU

IMPRINT

Automating Society
Taking Stock of Automated Decision-Making in the EU

1st edition, January 2019

Available online at www.algorithmwatch.org/automating-society

Automating Society –
Taking Stock of
Automated Decision-
Making in the EU

/ CONTENTS

KEY OF COLOURS:
EU countries covered in this report
EU countries not covered
non-EU countries
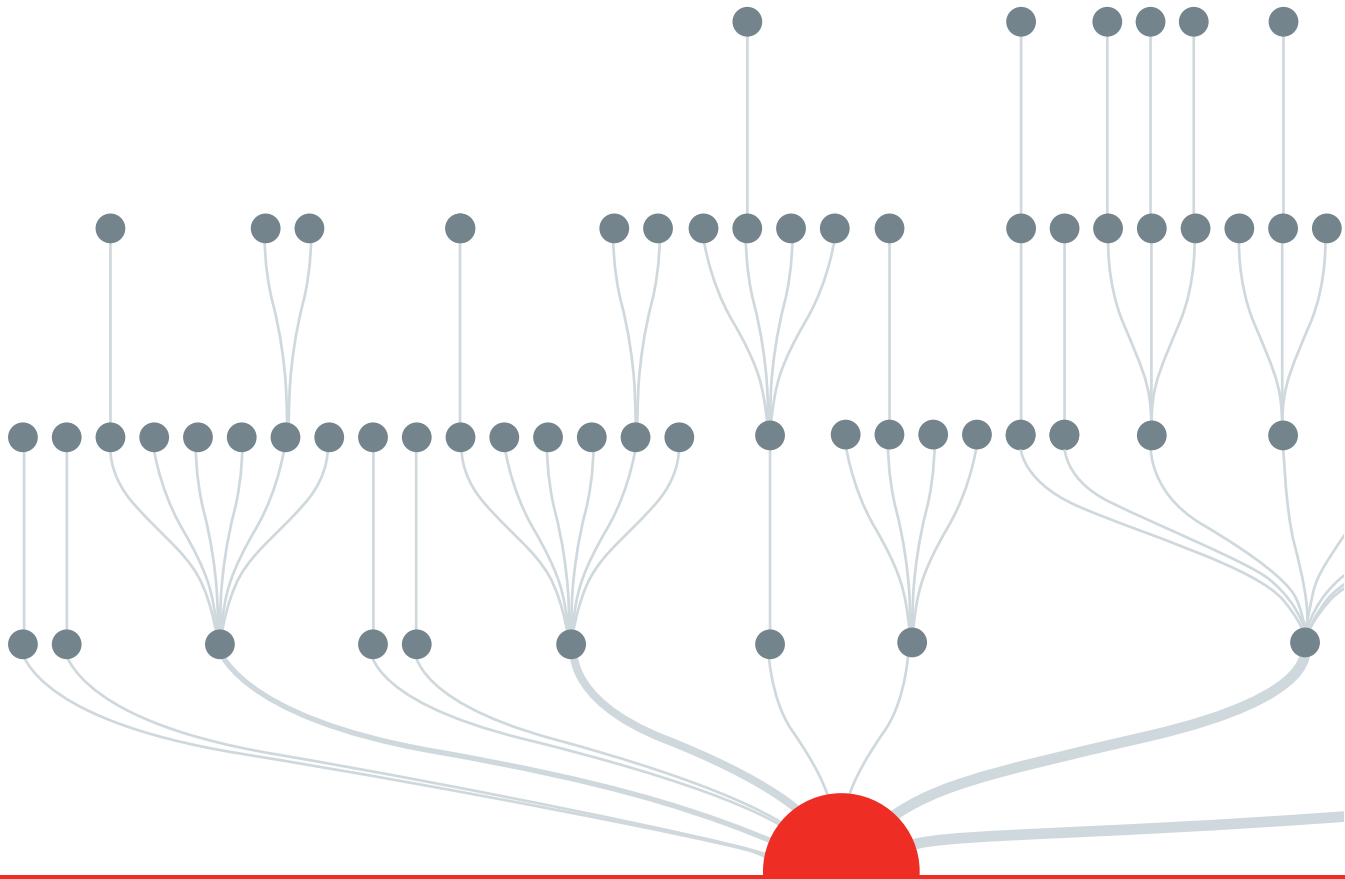
/ INTRODUCTION

Imagine you're looking for a job. The company you are applying to says you can have a much easier application process if you provide them with your username and password for your personal email account. They can then just scan all your emails and develop a personality profile based on the result. No need to waste time filling out a boring questionnaire and, because it's much harder to manipulate all your past emails than to try to give the 'correct' answers to a questionnaire, the results of the email scan will be much more accurate and truthful than any conventional personality profiling. Wouldn't that be great? Everyone wins—the company looking for new personnel, because they can recruit people on the basis of more accurate profiles, you, because you save time and effort and don't end up in a job you don't like, and the company offering the profiling service because they have a cool new business model.

When our colleagues in Finland told us that such a service actually exists, our jaws dropped. We didn't want to believe it, and it wasn't reassuring at all to hear the company claimed that basically no candidate ever declined to comply with such a request. And, of course, it is all
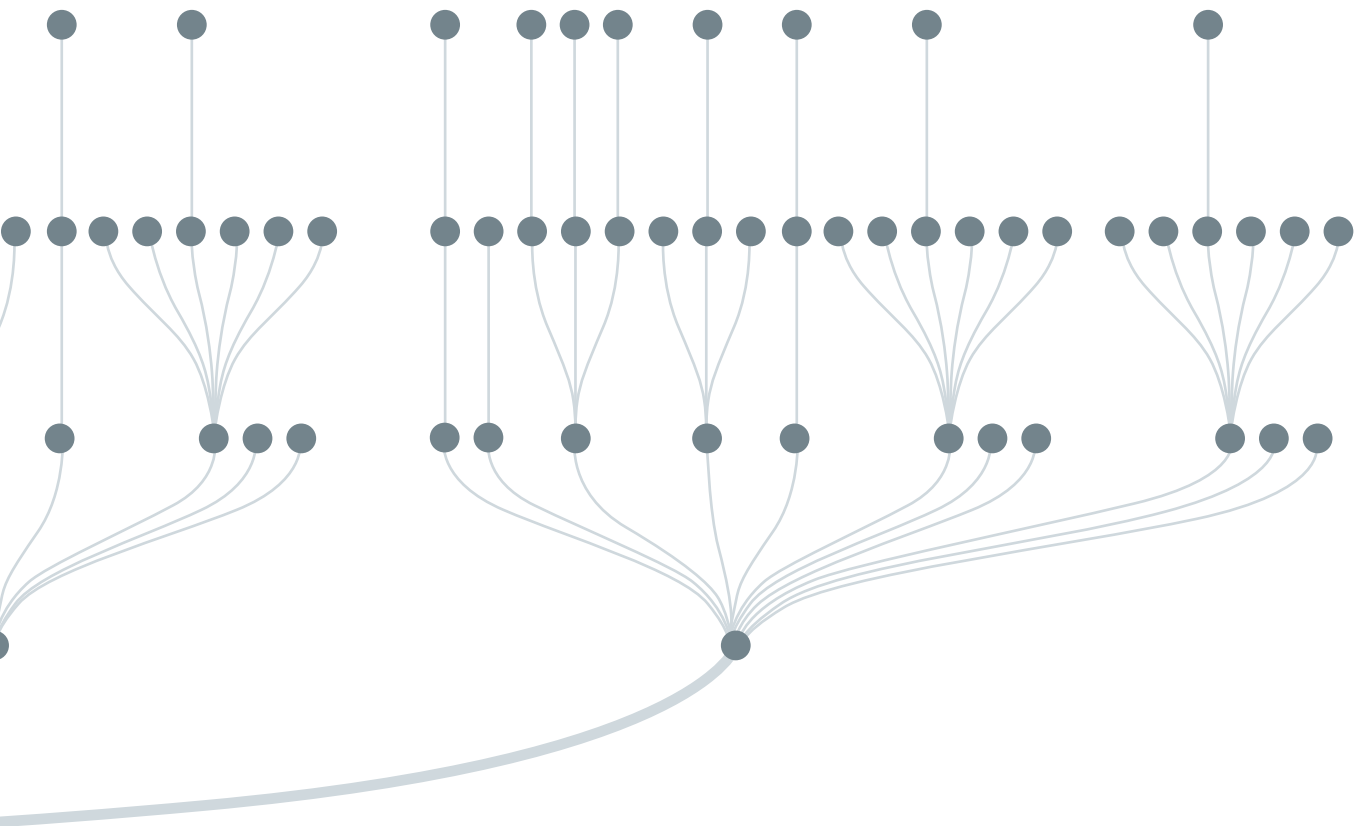
perfectly legal because job-seekers give their informed consent to open up their email to analysis—if you believe the company, that is. When we asked the Finnish Data Protection Ombudsman about it, he wasn't so sure. He informed us that his lawyers were still assessing the case, but that it would take a couple of weeks before he could give his opinion. Since this came to light just before we had to go to press with this publication, please go to our website to discover what his final assessment is.

/ Is automation a bad thing?

In many cases, automating manual processes is fantastic. Who wants to use a calculator instead of a spreadsheet when doing complex calculations on large sets of numbers (let alone a pen, paper and an abacus)? Who would like to manually filter their email for spam any more? And who would voluntarily switch back to searching for information on the Web by sifting through millions of entries in a Yahoo-style catalogue instead of just typing some words into Google's famous little box? And these examples do not even take into account

the myriads of automated processes that enable the infrastructure of our daily lives—from the routing of phone calls to automated electricity grid management, to the existence of the Internet itself. So we haven't just lived in a world of automated processes for a very long time; most of us (we'd argue: all of us) enjoy the many advantages that have come with it.

That automation has a much darker side to it has been known for a long time as well. When in 1957 IBM started to develop the Semi-Automated Business Research Environment (SABRE) as a system for booking seats on American Airlines' fleet of planes, we can safely assume that the key goal of the airline was to make the very cumbersome and error-prone manual reservation process of the times more effective for the company and more convenient for the customers. However, 26 years later, the system was used for very different purposes. In a 1983 hearing of the US Congress, Robert L. Crandall, president of American Airlines, was confronted with allegations of abusing the system—by then utilised by many more airlines—to manipulate the reservation process in order to favour American Airlines' flights over those of its competitors. His answer: "The preferential display of our flights, and the corresponding increase in our market share, is the competitive raison d'etre for having created the system in the first place".

"Why would you build and operate an expensive algorithm if you can't bias it in your favor?"

In their seminal paper "Auditing Algorithms", Christian Sandvig et al. famously proposed to call this perspective *Crandall's complaint:* "Why would you build and operate an expensive algorithm if you can't bias it in your favor?" ☑ **[IN 1]** In what is widely regarded as the first example of legal regulation of algorithmically controlled, automated decision-making systems, US Congress in 1984 passed a little-known regulation as their answer to Crandall's complaint. Entitled "Display of Information", it requires that each airline reservation system "shall provide to any person upon request the current criteria used in editing and ordering flights for the integrated displays and the weight given to each criterion and the specifications used by the system's programmers in constructing the algorithm."

/ A changed landscape

If a case like this sounds more familiar to more people today than it did in 1984, the reason is that we've seen more automated systems developed over the last 10 years than ever before. However, we have also seen more of these systems misused and criticised. Advances in data gathering, development of algorithms and computing power have enabled data analytics processes to spread out into fields so far unaffected by them. Sifting through personal emails for personality profiling is just one of the examples; from automated processing of traffic offences in France (see p. 70), allocating treatment for patients in the public health system in Italy (p. 88), automatically identifying which children are vulnerable to neglect in Denmark (p. 50), to predictive policing systems in many EU countries—the range of applications has broadened to almost all aspects of daily life.

Having said all this, we argue that there is an answer to the question 'Is automation a bad thing?' And this answer is: It depends. At first glance, this seems to be a highly unsatisfactory answer, especially to people who need to make decisions about the questions of how to deal with the development, like: Do we need new laws? Do we need new oversight institutions? Who do we fund to develop answers to the challenges ahead? Where should we invest? How do we enable citizens, patients, or employees to deal with this? And so forth. But everyone who has dealt with these questions already knows that it is the only honest answer.

/ Why *automated decision*-making instead of *Artificial Intelligence?*

One of the first hard questions to answer is that of defining the issue. We maintain that the term automated decision-making (ADM) better defines what we are faced with as societies than the term 'Artificial Intelligence', even though all the talk right now is about 'AI'.

Algorithmically controlled, automated decision-making or decision support systems are procedures in which decisions are initially—partially or completely—delegated to another person or corporate entity, who then in turn use automatically executed decision-making models to perform an action. This delegation—not of the decision itself, but of the execution—to a data-driven, algorithmically controlled system, is what needs our attention. In comparison, Artificial Intelligence is a fuzzily defined term that encompasses a wide range of controversial ideas and therefore is not very useful to address the issues at hand. In addition, the term 'intelligence' invokes connotations of a human-like autonomy and intentionality that should not be ascribed to machine-based procedures. Also, systems that would not be considered Artificial Intelligence by most of today's definitions, like simple rule-based analysis procedures, can still have a major impact on people's lives, i.e. in the form of scoring systems for risk assessment.

Algorithmically controlled, automated decision-making or decision support systems are procedures in which decisions are initially—partially or completely—delegated to another person or corporate entity, who then in turn use automatically executed decision-making models to perform an action.

In this report, we will focus on systems that affect justice, equality, participation and public welfare, either directly or indirectly. By saying *systems* instead of *technologies* we point to the fact that we need to take a holistic approach here: an ADM system, in our use of the term, is a socio-technological framework that encompasses a decision-making model, an algorithm that translates this model into computable code, the data this code uses as an input—either to 'learn' from it or to analyse it by applying the model—and the entire political and economic environment surrounding its use. This means that the decision itself to apply an ADM system for a certain purpose—as well as the way it is developed (i.e. by a public sector entity or a commercial company), procured and finally deployed—are parts of this framework.

Therefore, when an ADM system like SyRI is used in the Netherlands to detect welfare fraud (see p. 101), we not only need to ask what data it uses, but whether the use of this data is legal. We also need to ask what decision-making model is applied and whether it has a certain problematic bias, i.e. because it used a biased data set or was developed by people with underlying prejudices that were not controlled for. Other questions then arise: why did the government come up with the idea to use it in the first place? Is it because there is a problem that cannot be addressed in any other way, maybe due to its inherent complexity? Is it because austerity measures led to a situation where there are no longer enough case workers, so automation is used as an option to save money? Or is it because of a political decision to increase pressure on poor people to take on low-paying jobs?

/ The focus of this report

All these aspects need to be taken into account when asking whether automation can help solve a problem. This is why we decided to focus on four different issues in this report:

- **How is society discussing automated decision-making?** Here we look at the debates initiated by governments and legislators on the one hand, like AI strategies, parliamentary commissions and the like, while on the other hand we list civil society organisations that engage in the debate, outlining their positions with regard to ADM.

- **What regulatory proposals exist?** Here, we include the full range of possible governance measures, not just laws. So we ask whether there are ideas for self-regulation floating around, a code of conduct being developed, technical standards to address the issue, and of course whether there is legislation in place or proposed to deal with automated decision-making.

- **What oversight institutions and mechanisms are in place?** Oversight is seen as a key factor in the democratic control of automated decision-making systems. At the same time, many existing oversight bodies are still trying to work out what sectors and processes they are responsible for and how to approach the task. We looked for examples of those who took the initiative.

- Last but not least: **What ADM systems are already in use?** We call this section ADM in Action to highlight a lot of examples of automated decision-making already being used all around us. Here, we tried to make sure that we looked in all directions: do we see cases where automation poses more of a risk, or more of an opportunity? Is the system developed and used by the public sector, or by private companies?

/ The goals of this report

When we set out to produce this report, we had four main goals in mind:

1. To show that algorithmically driven, automated decision-making (ADM) systems are already in use all over the EU. So far, the discussion around the use of these systems, their benefits and risks, has been dominated by examples from the US: assessing the recidivism risk of criminals determining whether they are released on parole or stay in jail; teachers being fired based on their automatically calculated performance scores; people in minority neighbourhoods paying higher car insurance premiums than people from wealthy areas with the same risk. So we want to make clear what similar and other ADM systems are in use in the EU, in order to better inform the discussion about how to govern their use.

2. To outline the state of the political discussion not just on the EU level, but also in the member countries. We all know that Europe's diversity can be a burden when it comes to information flow across borders, especially because of 24 different official languages. So it was clear to us that we needed to change this situation as best we could by providing in-depth research from member countries in a shared language accessible to policy makers on the EU level. We approached this challenge in the best of the Union's traditions: As a cross-border, trans-disciplinary collaboration, pooling contributors from 12 different countries who speak their countries' language(s) and understand their societies' cultural contexts.

3. To serve as the nucleus for a network of researchers focusing on the impact of automated decision-making on individuals and society. This network includes journalists specialising in the nascent field of algorithmic accountability reporting, academics from economics, sociology, media studies, law and political sciences, to lawyers working in civil society organisations looking at the human rights implications of these developments. We will attempt to build on this first round of research and extend the network in the coming years because it is crucial to also include the many countries not covered in this initial report.

4. **To distil recommendations from the results of our findings:** for policy makers from the EU parliament and Member States' legislators, the EU Commission, national governments, researchers, civil society organisations (advocacy organisations, foundations, labour unions etc.), and the private sector (companies and business associations). You will find these recommendations on page 13.

/ The scope of this report

We view this report as an explorative study of automated decision-making both on the EU level and in selected Member States. It contains a wide range of issues and examples that justify a closer look, more in-depth research and discussion[1].

- Geography: For the initial edition of this report, we were not in a position to cover all 28 Member States. Instead, we decided to focus on a number of key countries, while making sure that the EU as a whole would be properly represented. Also, we deliberately excluded examples of applications coming from outside the EU. We all know that not only GAFA (Google, Amazon, Facebook and Apple) but also IBM, Microsoft, Salesforce and other US companies have a dominant market share in many sectors and also in European countries. Still, we confined ourselves to companies from Member States to better showcase their important role in automated decision-making.

- Topic: As laid out in the section "*Why automated decision-making instead of Artificial Intelligence?*", the definatory framework on which this report is based is still in an early phase. Therefore, much of the discussion between the contributors centred around the question whether a certain example fits the definition and should be part of the report or not. Most of the time, when there was latitude, we decided to err on the side of including it—because we see value in knowing about borderline cases at a stage when all of us are still trying to make sense of what is happening. When looking at examples framed as 'AI', 'big data' or 'digitisation/digitalisation'—like the plethora of AI and big data strategies, politically convened commissions and so forth—we focused on aspects of automated decision-making and decision support whenever they were referenced. In case automation as a term was not used, we still scanned for key concepts like discrimination, justice, equity/equality to take a closer look, since automation processes tend to be hiding behind these keywords.

Most of the time, when there was latitude, we decided to err on the side of including an example—because we see value in knowing about borderline cases at a stage when all of us are still trying to make sense of what is happening.

- Depth: Contributors to the report had approximately five days to research and complete their investigations. This means that they were not in a position to do any kind of investigative research, file freedom of information requests, or follow up resolutely in case companies or public entities denied their requests for information.

We would like to express our enormous gratitude to those who contributed to this report. As you can imagine, compiling it was a very special venture. It required a lot of expertise, but also spontaneity, flexibility and determination to piece this puzzle together. As always, this also means it required, at times, quite a lot of tolerance and patience when deadlines neared and pressure mounted. We consider ourselves lucky to have found all this in our

---

1   If funding permits, we would like to extend this exploration to all EU countries in the future. So if there is anything you know of that you think should have been part of this report, please tell us by writing to eu-adm-report@algorithmwatch.org.

colleagues. You can learn who they are in more detail in the information boxes underneath the country chapters.

We are also greatly indebted to Becky Hogge and the entire team at OSF's information programme. When we approached Becky with the idea to produce this report, she was not only immediately willing to take the risk of embarking on this journey but she also saw the application through in almost no time. As a result, we were able to seize this opportunity to provide valuable input at a decisive stage of the debate.

/ Many shades of grey

Offering to create a profile by scanning your email seems to be a clear-cut case of a misuse of power. But is it?

In closing, let's come back to the case of the company offering to create a profile by scanning your email. It seems to be a clear-cut case of a misuse of power. If someone is looking for a job, how can they be in a position to freely give their consent to a procedure the company filling a position wants them to undergo—especially if this is something as private as analysing their personal emails? This seems to be a highly unethical, if not illegal, proposal that we cannot possibly accept.

Now consider the following scenario: The reliability of the tests that recruiters ask candidates to take is highly controversial due to their social desirability bias, meaning that people tend to give answers not in the most truthful manner but in a manner that they think will be viewed favourably by others. The company argues that this bias can be minimised by looking at a corpus of emails that is very hard to manipulate. Imagine that this claim is supported by sound research. In addition, the company argues that candidates' emails are never permanently stored anywhere, they are only passed through the company's system in order to apply its analysis procedure to them and then disappear. Also, no human will ever look at any individual email. Now imagine this procedure is audited by a trusted third party—like a data protection authority, an ombudsman, a watchdog organisation—who confirms the company's claim. Lastly, imagine that all candidates are provided with information about this process and the logic of the underlying profiling model, and then given the choice whether to undergo the 'traditional' process of filling out questionnaires, or to allow the company to scan their emails.

Given all these assumptions, would this change your mind regarding the question whether we can approve of this procedure, whether we should legally permit this procedure, and probably even whether you personally might consent to such a procedure? Mind you, we're not going to give away what our decision is. Instead, we'll leave you to contemplate the countless shades of grey in a world of automated decision-making and hope that you are inspired by reading this report.

Brigitte Alfter, Ralph Müller-Eiselt, Matthias Spielkamp

/ RECOMMENDATIONS
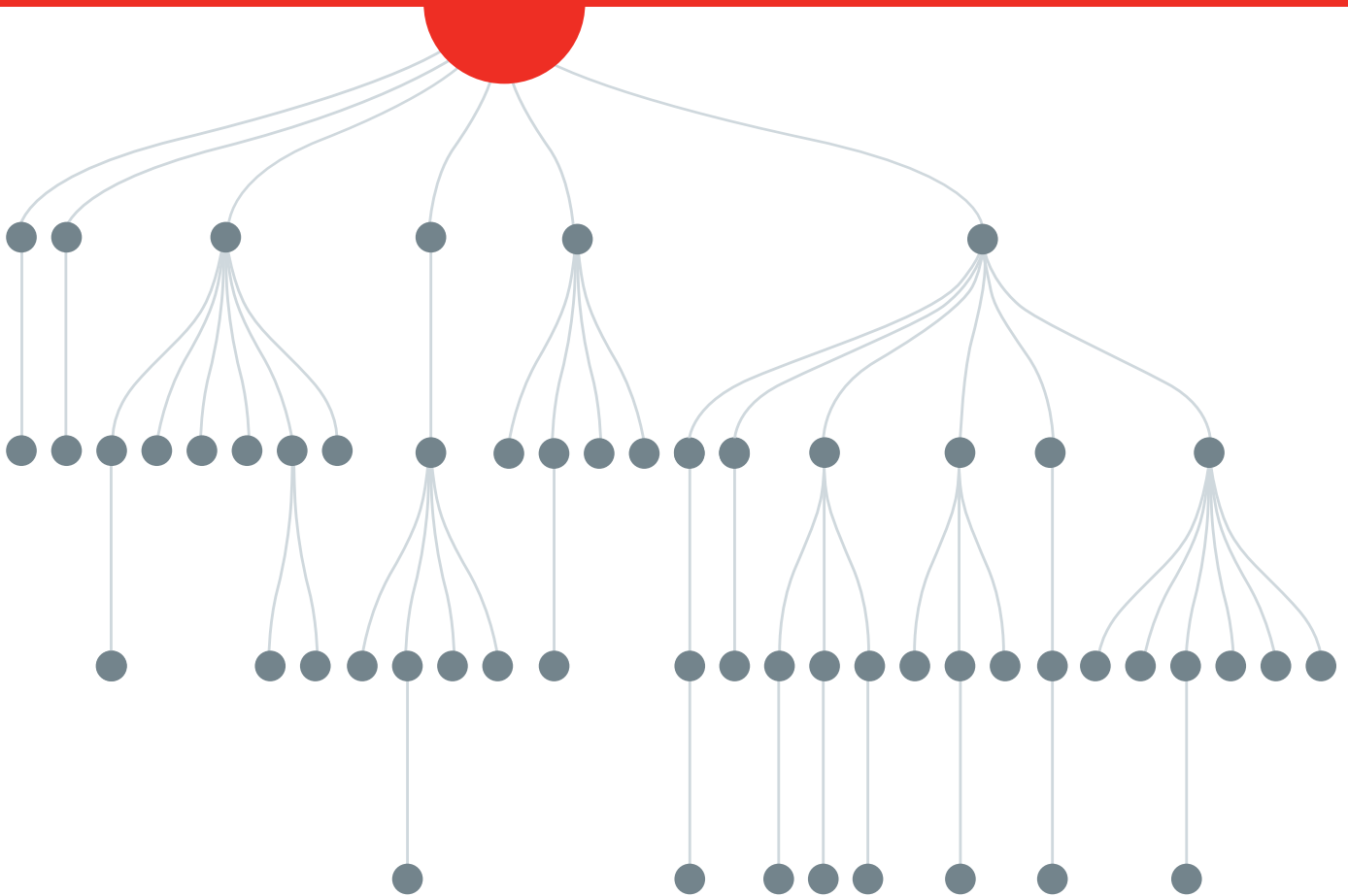
In this report we have compiled findings from 12 EU member states and the level of the EU. In all countries, we looked at the debates focused on the development and application of automated decision-making systems. We identified regulatory strategies and compiled examples of ADM systems already in use. Based upon these findings, we have the following recommendations for policy makers in the EU parliament and Member States parliaments, the EU Commission, national governments, researchers, civil society organisations (advocacy organisations, foundations, labour unions etc.), and the private sector (companies and business associations). These recommendations are not listed in order of importance and we have refrained from referring to specific examples.

Focus the discussion on the politically relevant aspects. 'Artificial Intelligence' is all the rage right now, ranging from debates about relatively simple rule-based analysis procedures to the threat of machine-created 'super-intelligence' to humanity. It is crucial to understand what the current challenges to our societies are. 'Predictive analytics', used to

forecast maintenance issues on production lines for yoghurt, should not concern us too much—except maybe when it relates to the allocation of research grants. However, predictive analytics used for forecasting human behaviour, be it in elections, criminal activity, or of minors, is an entirely different story. Here, we need to guarantee that these systems are identified as crucial for society. They must be democratically controlled by a combination of regulatory tools, oversight mechanisms, and technology.

The debate around AI should be confined to current or imminent developments. There is a lot of discussion around the ideas of 'artificial general intelligence', 'super-intelligence', 'singularity' and the like. As attractive as these discussions may appear to some, right now they are based on mere speculation and therefore they are a distraction from the most pressing question: how to deal with current developments?

Consider automated decision-making systems as a whole, not just the technology. Automated decision-making processes are often framed as technology. As a result, the debate revolves around questions of accuracy, data quality and the like. This risks overlooking many of the crucial aspects of automated decision-making: the decision itself to apply an ADM system for a certain purpose, the way it is developed (i.e. by a public sector entity or a commercial company), and how it is procured and finally deployed. These are all part of the framework that needs to be reflected when discussing the pros and cons of using a specific ADM application. This means that we should ask what data the system uses, whether the use of this data is legal, what decision-making model is applied and whether it has a problematic bias. But we also need to ask why companies or governments come up with the idea to use specific ADM systems in the first place. Is it because of a problem that cannot be addressed in any other way, maybe due to the inherent complexities associated with the problem? Have austerity measures led to a situation where there are not enough resources for humans to deal with certain tasks, so automation is used as an option to save money?

Empower citizens to adapt to new challenges. There are a number of ways we can enhance citizens' expertise to enable them to better assess the consequences of automated decision-making. We would like to highlight the Finnish example: in order to support the goal of helping Finns understand the challenges ahead, the English-language online course Elements of Artificial Intelligence was developed as a private-public partnership and is now an integral part of the Finnish AI programme. This freely available course teaches citizens about basic concepts and applications of AI and machine learning. Almost 100,000 Finns have enrolled in the course to date, thus increasing public understanding of AI and enabling them to participate in public debate on the subject. On the course, some societal implications of AI are introduced, for example, algorithmic bias and de-anonymisation, to underscore the need for policies and regulations that help society to adapt more easily to the use of AI.

Empower public administration to adapt to new challenges. Public administration not only procures a lot of automated decision-making systems, it also uses them for purposes that have a big impact on individuals and society, i.e. border control, crime prevention and welfare management. Public administration must ensure a high level of expertise inside its own institutions in order to either develop systems themselves, or at least be in a position to meaningfully oversee outsourced development. This can be achieved by creating public research institutions, i.e. in cooperation with universities or public research centres, that can teach and advise civil servants. Such institutions should also be created at the EU level to assist Member States.

Strengthen civil society involvement. Our research shows that, even in some large Member States, there is a lack of civil society engagement and expertise in the field of ADM. This shortcoming can be addressed by a) civil society organisations identifying the consequences of automated decision-making as a relevant policy field in their countries and strategies, b) grant-making organisations earmarking parts of their budget for ADM, developing funding calls, and facilitating networking opportunities, c) governments making public funds available to civil society interventions.

Make sure adequate oversight bodies exist and are up to the task. Oversight over automated decision-making systems is organised by sector. For example, there are different oversight bodies for traffic (automated cars), finance (algorithmic trading), and banking (credit scoring). This makes sense, since many systems need to be looked at in their respective contexts, with specific knowledge. It is doubtful, though, that many of the oversight bodies in place have the expertise to analyse and probe modern automated decision-making systems and their underlying models for risk of bias, undue discrimination, and the like. Here, Member States and the EU are called upon to invest in applied research to enable existing institutions to catch up, or to create new ones where needed. In addition, parliaments and courts need to understand the fact that it is they who need to oversee the use of ADM systems by governments and public administrations. Therefore, they also need appropriate assistance in order to empower them to live up to the task.

Close the gap between Member States. Many countries in the EU have developed strategies for 'digitisation/digitalisation', 'big data', or 'Artificial Intelligence'. Still, some countries are lagging behind when it comes to discussing the consequences that automated decision-making can have on individuals and society—either because of a lack of resources, or because of differing priorities. Since the question of whether and how ADM systems should be used very much depends on the cultural context, expertise in Member States is needed. Some Member States should invest more in capacity building. The EU is doing a lot in this regard, i.e. by developing its own recommendations via the EU High-Level Expert Group on AI. In addition, Member States can use this report to get an idea of what is going on in other countries to see whether there is a need to catch up. If that is the case they can use this as an opportunity to create structures and mechanisms that help decision makers and the general public learn from leading countries. On the national level, this could happen in the form of intergovernmental working groups or bi- and multi-national cooperation. At the EU level, the EU Commission should continue to offer forums like the European AI Alliance to further this goal.

Don't just look at data protection for regulatory ideas. Article 22 of the General Data Protection Regulation (GDPR) has been the focus of many discussions about automated decision-making in recent years. A consensus is starting to develop that a) the reach of Article 22 is rather limited (see page 42) and b) that there are use cases of ADM systems that cannot be regulated by data protection (alone), i.e. predictive policing. These systems can have effects on communities – like over-policing – without the use of personal data, so the GDPR doesn't even apply. At the same time, since no individual is discriminated against if a neighbourhood slowly turns into a highly patrolled area, anti-discrimination laws are of no help either. There needs to be a discussion about how to develop governance tools for these cases. In addition, stakeholders need to look at creative applications of existing regulatory frameworks, like equal-pay regulation, to address new challenges like algorithmically controlled platform work, also known as the Gig Economy, and explore new avenues for regulating the collective effects of ADM altogether.

Involve a wide range of stakeholders in the development of criteria for good design processes and audits, including civil liberty organisations. In some of the countries we surveyed, governments claim that their strategies involve civil society stakeholders in the current discussion in order to bring diverse voices to the table. However, it is often the case that the term civil society is not well defined. It can be legitimately argued that the term civil society also includes academia, groups of computer scientists or lawyers, think tanks and the like. But if governments use this broad definition to argue that 'civil society' is on board, yet civil liberty organisations are not part of the conversation, then very important viewpoints might be missed. As a result, there is a risk that the multi-stakeholder label turns out to be mere window dressing. Therefore, it is critical that organisations focused on rights are included in the debate.

The EU funded DANTE and TENSOR projects plan to offer law enforcement agencies a platform for „planning and prevention applications for early detection of terrorist activities, radicalisation and recruitment" and aim to offer far reaching "automated functionalities".

25 EU countries signed the Declaration of Cooperation on Artificial Intelligence. With regard to automated decision-making processes, the signatories commit to ensure "that humans remain at the centre of the development, deployment and decision-making of AI".

BELFAST

Latavia

iBorderCtrl is a system tested in Hungary, Greece and Latvia to screen non-EU nationals at EU borders, using automated interviews with a virtual border guard, based on "deception detection" technology".

The EU AI Strategy states that "AI systems should be developed in a manner which allows humans to understand (the basis of) their actions".

BRUSSELS

LUXEMBOURG

Hungary

Serbia

A general framework for equal treatment in employment and occupation has been established in the EU. It not only includes formally employed persons, as many national regulations do, but could be applicable to platform workers and their automated management.

Maze-donia

Greece

EUROPEAN UNION

BY KRISTINA PENNER

The EU is a lot of things: It is a union of 28 Member States that has a commission, a parliament, a council of national ministers, the Court of Justice, and a couple of other institutions. It has many funding programmes for research and development, committees such as the European Economic and Social Committee (EESC), an ombudsman and agencies like one for fundamental rights. In addition, there are scores of business and civil society pressure groups, professional societies and think tanks that try to interact with the EU institutions or influence policy making.

So when we started looking into where automated decision-making (ADM) is currently being discussed in the EU and what regulatory proposals already exist, we had many directions in which to look. Some places were obvious: The EU's Declaration of cooperation on Artificial Intelligence (AI), the Commission's Communication *Artificial Intelligence for Europe*, its High-Level Expert Group focused on AI, the European Parliament's resolution on robotics and the European Economic and the Social Committee's opinion on Artificial Intelligence (AI). These all state that the aim of using ADM is to simultaneously maximise the benefit to society, help business, spur innovation and encourage competition.

The term Artificial Intelligence is commonly used in discussions, however upon closer inspection it becomes clear that it is algorithmically controlled ADM systems that are at the centre of many of these debates. The EU's algo:aware project is a case in point. Its mandate is to analyse "opportunities and challenges emerging where algorithmic decisions have a significant bearing on citizens, where they produce societal or economic effects which need public attention."[1]

This perspective should be welcomed, as it focuses on a key question: How much of our autonomy are we willing to cede to automated systems? And it begs another question: Do we have the appropriate safeguards in place to help us understand what we're doing and enable us to stay in control of the process?

One of these safeguards, the General Data Protection Regulation (GDPR), has been hailed as one of the most powerful tools the EU has come up with to address automated decision-making. However, many experts now agree that it has shortcomings; others fear that, when it comes to tackling the challenges posed by ADM systems, the GDPR may not be of any help at all. Therefore, it could well be time to look in a different direction to discover other means that can be used to deal with these challenges. EU employment equality law could be one such direction. Alternatively, long-existing regulations, such as the Financial Market Directive with its rules for high frequency algorithmic trading, might serve as a model. Last but not least, civil society and standards bodies have begun to play an active role in shaping the discussion.

---

1   The project's first version of its so-called "state-of-the-art report on algorithmic decision-making" was published too late for consideration in this publication—it is available at https://www.algoaware.eu/wp-content/uploads/2018/08/AlgoAware-State-of-the-Art-Report.pdf

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – EUROPEAN INSTITUTIONS

/ Declaration of cooperation on Artificial Intelligence (AI)

On April 10, 2018, twenty-five European countries signed the Declaration of Cooperation on Artificial Intelligence (AI) ⧉ **[EU 1]** with the stated goal to build on "the achievements and investments of Europe in AI" ⧉ **[EU 1]**, as well as progress towards the creation of a Digital Single Market. In the declaration, participating Member States agree to shape a European approach to AI, increase public and private investment, and commit to publish a coordinated plan on AI before the end of 2018. While the declaration focuses on increasing "competitiveness, attractiveness and excellence in R&D in AI", it also states that the signatories want to foster the "development of AI in a manner that maximizes its benefit to economy and society" and "exchange views on ethical and legal frameworks related to AI in order to ensure responsible AI deployment." With regard to automated decision-making processes, the signatories commit to "ensure that humans remain at the centre of the development, deployment and decision-making of AI, prevent the harmful creation and use of AI applications, and advance public understanding of AI". In addition, the accountability of AI systems is supposed to be increased.

/ Communication on Artificial Intelligence for Europe

Two weeks after the publication of the Declaration of cooperation on Artificial Intelligence, on April 25, 2018, the European Commission published a Communication on Artificial Intelligence for Europe. ⧉ **[EU 2]** Following the previously adopted declaration, it elaborates on the three pillars outlined as the core of the proposed strategy:

1. Being ahead of technological developments and industrial capacity and strengthening public and private actors, including not only investments in research centres, but also the development of an "AI-on-demand platform" to provide data and resources for the creation of a data economy
2. To increase preparedness for socio-economic changes—by modernising education and training systems—supporting labour market transitions
3. To develop AI ethics guidelines and to ensure legal clarity in AI based applications (announced for 2019)

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

Although the Communication does not differentiate between AI and other systems of ADM in some parts of its strategy outline, clear indications of responses to ADM can be found in specific measures and policy packages introduced by the Communication.

Preparing for socio-economic changes

The Communication states that its goal is "to lead the way in developing and using AI for good and for all" and to follow an approach that "benefits people and society as a whole". The main steps and measures outlined in the initiative focus on competitiveness and investments. It also pledges to encourage diversity and interdisciplinarity ("diversity and gender balance in AI jobs"). The Communication explicitly addresses the risk posed by automated decision-making: "Some AI applications may raise new ethical and legal questions, related to liability or fairness of decision-making". These risks are to be mitigated mainly by AI ethics guidelines (see below in this chapter) and by providing guidance on the interpretation of the Product Liability Directive.

Research and innovation on responsible and explainable AI

> "AI systems should be developed in a manner which allows humans to understand (the basis of) their actions. [...] Whilst AI clearly generates new opportunities, it also poses challenges and risks, for example in the areas of safety and liability, security (criminal use or attacks), bias and discrimination." ☑ **[EU 2]**

The Commission announced that it will support (basic and industrial) research and innovation in fields built on the guiding principle of "responsible AI", including investment and encouragement of research and testing in sectors such as health, security, public administration and justice, with the goal to enable policy makers to gain experience and to devise suitable legal frameworks.

Research on the development of explainable AI systems—and unsupervised machine learning in order to increase transparency and minimise the risk of bias and errors—is only planned beyond 2020.

One starting point is a pilot project[2] commissioned by the EU Commission on algorithmic awareness building. Recognising that algorithms play an increasingly important role in decisions of relevance to public policy, the Commission procured an in-depth analysis into algorithmic transparency. This aims to raise awareness and build an evidence base for the challenges and opportunities of algorithmic decisions. The project's objective, among others, is to design or prototype solutions for a selection of problems. These include policy responses, technical solutions and private sector and civil society-driven actions in response to the challenges brought by ADM, including bias and discrimination.

Building a European Data Economy and creating a European Artificial Intelligence-on-demand-platform

Already embedded in the Digital Single Market strategy—translated into proposals for action in its mid-term review, and now applied to AI—the EU Commission is striving to establish and continue the expansion of a "common European data space". ☑ **[EU 4]** Identifying data as a key ingredient for competitive AI, the EU plans "a seamless digital area with the scale to enable the development of new products and services based on data". This is to be realised, in particular, by increasing the accessibility and re-usability of public sector information, the proposal to amend the Public Sector Information (PSI) Directive being a core element. ☑ **[EU 5]** The second comprehensive initiative in this regard is the creation of a European AI-on-demand platform aiming to strengthen a European AI community, which is already taking shape as AI4EU ☑ **[EU 6]** (see below).

The European Commission aims to facilitate the re-use of PSI such as legal, traffic, meteorological, economic and financial data throughout the European Union. This will be done by harmonising the basic conditions that make PSI available to re-users, to enhance the development of Community products and services based on PSI, and to avoid distortions in competition. Stakeholder concerns are especially related to the necessary protection of personal data, especially in sensitive sectors such as health, when they take the decision on the re-use of PSI (see paragraph on the EDPS).

☑ **LINKS:** *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

---

2   The project algo:aware ☑ **[EU 3]** aims to engage with a range of stakeholders and seeks to map the areas of interest where algorithmic operations bear significant policy implications. So far, they have produced a series of reports, blog posts, case studies, infographics and policy developments. They aim to provide a platform for information, and a forum for informed debate on the opportunities and challenges that algorithmic decision-making can provide in commercial, cultural, and civic society settings.

The development of ethics guidelines

The Communication—and the proposed implementation process behind the EU Initiative on AI—addresses ethical and legal questions. It states that the basis for work on these questions will be the EU's values laid down in the Charter of Fundamental Rights of the EU.

By the summer of 2019, the Communication foresees the creation of a framework for all relevant stakeholders and experts—see European AI Alliance and High-Level Expert Group in this chapter—who will draft the guidelines, addressing issues such as future of work, fairness, safety, security, social inclusion and algorithmic transparency. More broadly, the group will look at the impact on fundamental rights, including privacy, dignity, consumer protection and non-discrimination.

The development of the AI Ethical Guidelines will build on the Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems by the European Group on Ethics in Science and New Technologies (EGE) ⧉ **[EU 7]**[3]. They will tackle questions on "liability or potentially biased decision-making". It affirms that the EU must develop and apply a framework to protect ethical principles such as accountability and transparency, to reach the indicated goal to become "the champion of an approach to AI that benefits people and society as a whole". ⧉ **[EU 2]**

The development of the draft is accompanied by assessments from the EU Fundamental Rights Agency (FRA, see below), and the evaluations of the EU Safety framework.

The Commission is planning to systematically monitor AI-related developments, including policy initiatives in Member States in order to develop an AI index to inform the discussion.

By announcing work on a coordinated plan with Member States by the end of the year—also strongly focusing on competition and economic aspects—there is a risk that many of these newly created bodies and consultation processes will be insufficiently incorporated into the implementation framework.

The following statements and resolutions are also referred to or announced in the Communication, but not included in this report:

- Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems of the European Group on Ethics in Science and New Technologies (EGE) ⧉ **[EU 7]**
- Communication on the future of connected and automated mobility in Europe ⧉ **[EU 8]**
- A renewed European Agenda for Research and Innovation—Europe's chance to shape its future ⧉ **[EU 9]**

/ European Economic and Social Committee opinion on AI

The European Economic and Social Committee (EESC) is a consultative body of the European Union and an advisory assembly (Article 300 TFEU). It is composed of "social partners" and economic and social interest groups—namely: employers/employers' organisations, employees/trade unions and representatives of various other interests.[4] ⧉ **[EU 10]**

---

3   The EGE is an independent advisory body of the President of the European Commission.

4   The Committee was set up by the 1957 Rome Treaties in order to involve economic and social interest groups in the establishment of the common market and to provide institutional machinery for briefing the Commission and the Council of Ministers on European issues.

In May 2017, the EESC adopted a so-called "own-initiative opinion" on the "consequences of artificial intelligence on the (digital) single market, production, consumption, employment and society" that was taken into account in the EU Communication on AI. In comparison to other papers and communiqués on the EU AI initiative, it presents a very precise, specific and in-depth analysis of different types and subfields of AI (narrow/general AI) and its parameters and consequences. The committee provides general recommendations, including a human-in-command approach for "responsible AI". It identifies eleven areas where AI poses societal and complex policy challenges, namely: ethics, safety, privacy, transparency and accountability, work, education and skills, (in-)equality and inclusiveness, law and regulation, governance and democracy, warfare and super-intelligence. ⧉ **[EU 11]**

In the paragraph on "Transparency, comprehensibility, monitorability and accountability", the document deals with actions and decisions of AI systems (through algorithms) in people's lives. The EESC argues that the acceptance, sustainable development and application of AI is based on the ability to understand, monitor and certify the operation, actions and decisions of AI systems, including retrospectively. The committee therefore advocates for transparent, comprehensible and monitorable AI systems whose operations have to be accountable, including retrospectively. In addition, the EESC demands that recommendations be made to determine which decision-making procedures can and cannot be transferred to AI systems, and when human intervention is desirable or mandatory.

The opinion offers a critical and substantial analysis of ethical questions, like embedded bias in AI development. It highlights the responsibility of humans to ensure that accuracy, data quality, diversity and self-reflection are taken into account in the design of AI, and to reflect on the environment in which it is applied. The EESC calls for a "code of ethics for the development, application and use of AI so that throughout their entire operational process AI systems remain compatible with the principles of human dignity, integrity, freedom, privacy and cultural and gender diversity, as well as with fundamental human rights". The committee also asks for the "development of a standardisation system for verifying, validating and monitoring AI systems" on a supra-national level.

/ EP Report on Civil Law Rules on Robotics / EP Committee on Robotics

In January 2017, the European Parliament adopted a report with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103/(INL) ⧉ **[EU 12]**). The report covers a wide range of different areas, such as liability rules, ethical principles, standardisation and safety, data protection, human enhancement or education and employment. IT also provides a Code of Ethical Conduct for Robotics Engineers and a Code for Research Ethics Committees.

The European Commission was asked by the Committee on Robotics to use its findings and recommendations as guidelines for the implementation of the EU AI strategy. This included the creation of a legal framework to address the use of AI and robotics for civil use. It specifically mentions that the "further development and increased use of automated and algorithmic decision-making undoubtedly has an impact on the choices that a private person (such as a business or an internet user) and an administrative, judicial or other public authority take in rendering their final decision of a consumer, business or authoritative nature", therefore "safeguards and the possibility of human control and verification need to be built into the process of automated and algorithmic decision-making", including "the right to obtain an explanation of a decision based on automated processing". ⧉ **[EU 12]**

In this context, a public consultation with an emphasis on civil law rules was conducted by the Parliament's Committee on Legal Affairs (JURI) to seek views on how to address the challenging ethical, economic, legal and social issues related to robotics and AI developments. ⬀ **[EU 13]**

The consultation did not contain questions about automated or algorithmic decision-making, neither in the general public nor in the expert version of the questionnaire. However, some submissions explicitly referred to "algorithmic discrimination" and "transparency of algorithmic decision-making". In addition, there was a mention that fundamental rights "would be at risk if unethical practice is facilitated by virtue of algorithms focused on commercial gain, for example because humans 'allow' or 'rely' on robot sorting techniques that are discriminatory and may be unfair and undermine dignity and justice". Another respondent wrote: "While the EU has started to address and regulate in the EU General Data Protection Regulation, the use of automated individual decision-making systems, such as profiling, further work is needed to fully understand the functioning of these systems and [to] develop adequate safeguards to protect human rights and dignity."

A highly controversial proposal appears in paragraph 59f of the resolution. It calls for the creation of "a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently". A group of around 285 "political leaders, AI/robotics researchers and industry leaders, physical and mental health specialists, law and ethics experts" signed an open letter on "Artificial Intelligence and Robotics" ⬀ **[EU 14]**, criticising the idea as misguided from a technical, ethical and legal perspective. Both the EESC's opinion (see above) in its article 3.33 ⬀ **[EU 15]** and UNESCO's COMEST report on Robotics Ethics of 2017 in its article 201 ⬀ **[EU 16]** state that they were opposed to any form of legal status for robots or AI.

/ European Union Agency for Fundamental Rights

The European Union Agency for Fundamental Rights (FRA ⬀ **[EU 17]**) is the EU's centre of fundamental rights expertise. Established in 2007 and based in Vienna, it is one of the EU's decentralised agencies set up to provide expert advice to EU institutions and Member States. The FRA is an independent EU body, funded by the Union's budget, and is a member of the EU High-Level Expert Group on AI.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

Stating that the intersection of rights and technological developments warrants a closer analysis, the FRA actively examines two aspects of automated—or data-driven, as they call it—decision-making: Its effects on fundamental rights and the potential for discrimination in using big data for ADM. In 2018, the FRA started a new project on "Artificial Intelligence, Big Data and Fundamental Rights", with the aim of contributing to the creation of guidelines and recommendations in these fields. ⬀ **[EU 18]** In a focus paper they point out that "When algorithms are used for decision-making, there is potential for discrimination against individuals. The principle of non-discrimination, as enshrined in Article 21 of the Charter of Fundamental Rights of the European Union (EU), needs to be taken into account when applying algorithms to everyday life." ⬀ **[EU 19]** It also suggests potential ways of minimising this risk, like a strong inter-disciplinary approach: "Although data protection principles provide some guidance on the use of algorithms, there is more that needs to be considered. This requires strong collaboration between statisticians, lawyers, social scientists, com-

puter scientists and subject area experts. In this way, a truly fundamental rights-compliant approach can be developed."

The project is collecting data on the fundamental rights implications and opportunities related to AI and Big Data in order to support development and implementation of policies.

The agency will further analyse the feasibility of carrying out online experiments and simulation case studies using modern data analysis tools and techniques.

In December 2018, the FRA published an update ⤢ [EU 20] of their "guide on preventing unlawful profiling" from 2010. The update takes into account legal and technological developments and the increased use of profiling by law enforcement authorities. The guide also widened its scope to include border management, and it offered "practical guidance on how to avoid unlawful profiling in police and border management operations."

/ AI4EU – European AI-on-demand platform

The main elements of the AI4EU project[5] ⤢ [EU 22] are the creation of a European AI-on-demand platform and the strengthening of the "European AI community". Other areas of action are called "Society and European Values", "Business and Economy", and "AI Research and Innovation". The platform is supposed to "act as a broker, developer and one-stop shop providing and showcasing services, expertise, algorithms, software frameworks, development tools, components, modules, data, computing resources, prototyping functions and access to funding." Among a long list of activities to reach this goal is the development of a "Strategic Research Innovation Agenda for Europe", including ethical, legal and socio-economic aspects.

With regard to ADM, it remains to be seen what the establishment of the AI4EU Ethical Observatory will look like. Its mandate is to ensure the respect of human centred AI values and European regulations.

/ High-Level Expert Group on Artificial Intelligence

To support the implementation of the European strategy on AI, the Commission appointed 52 experts, including representatives from academia, civil society and industry to the High-Level Expert Group on AI (AI HLEG). ⤢ [EU 23]

The group is tasked to prepare ethics guidelines[6] that will build on the work of the European Group on Ethics in Science and New Technologies, and of the European Union Agency for Fundamental Rights (see both in this chapter), published as the group's first deliverable in December2018. The guidelines will cover issues such as fairness, safety, transparency, the future of work, democracy and more broadly the impact of AI and automated decision-making on the application of the Charter of Fundamental Rights, including: privacy and personal data protection, dignity, consumer protection and non-discrimination.

---

5   The project website was updated just after the editorial deadline. The launch date is planned for 1 January 2019. ⤢ [EU 21]

6   The group's first draft of the guidelines were published too late for consideration in this publication; it is available at https://ec.europa.eu/futurium/en/system/files/ged/ai_hleg_draft_ethics_guidelines_18_december.pdf

The group is further mandated to support the Commission through recommendations regarding the next steps on how to address mid-term and long-term opportunities and challenges related to Artificial Intelligence. The AI HLEG's recommendations will feed into the policy development process, the legislative evaluation process, and the development of a next-generation digital strategy.

Overall, the AI HLEG serves as the steering group for the European AI Alliance's work (see below), and it interacts with other initiatives, helps stimulate a multi-stakeholder dialogue, gathers participants' views and reflects them in its analysis and reports. With these results it supports the Commission on further engagement and outreach mechanisms to interact with a broader set of stakeholders in the context of the AI Alliance.

/ European AI Alliance

The European AI Alliance is hosted and facilitated by the EU Commission. ⤴ **[EU 24]** Its members, including businesses, consumer organisations, trade unions, and other representatives of civil society bodies, are supposed to analyse the emerging challenges of AI, interact with the experts of the High-Level Expert Group on Artificial Intelligence (AI HLEG) in a forum-style setting, and to feed into the stakeholder dialogue. ⤴ **[EU 25]** By signing up to the Alliance, members get access to a platform where they can offer input and feedback to the AI HLEG. ⤴ **[EU 26]** The AI HLEG drew on this input when preparing its draft AI ethics guidelines and completing its other work. Moreover, the discussions hosted on the platform are supposed to directly contribute to the European debate on AI and to feed into the European Commission's policy-making in this field.

> "[…] Given the scale of the challenge associated with AI, the full mobilisation of a diverse set of participants, including businesses, consumer organisations, trade unions, and other representatives of civil society bodies is essential." ⤴ **[EU 25]**

The Alliance is also a place to share best practices, network and encourage activities related to the development of AI and is open to anyone who would like to participate in the debate on AI in Europe.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY

/ AccessNow

AccessNow is an international non-profit group focusing on human rights, public policy and advocacy. ⤴ **[EU 27]** In their new report on Human Rights in the Age of Artificial Intelligence ⤴ **[EU 28]**, they contribute to the analysis of AI and ADM and conceptualise its risks from a human rights perspective.

As human rights are universal and binding and more clearly defined than ethical principles, AccessNow assesses how human rights complement existing ethics initiatives, as "human rights can provide well-developed frameworks for accountability and remedy."

In their report, they examine how current and near-future uses of AI could implicate and interfere with human rights. They emphasise that the scale at which AI can identify, classify, and discriminate among people magnifies the potential for human rights abuses in both

reach and scope. The paper explores how AI-related human rights disproportionately harm marginalised populations. The paper also develops recommendations on how to address AI-related human rights harm.

/ BEUC – The European Consumer Organisation

BEUC (Bureau Européen des Unions de Consommateurs) is a Brussels umbrella group for European consumer protection organisations. It represents consumer organisations and defends interests of consumers at the EU level. BEUC investigates EU decisions and developments likely to affect consumers, in areas including AI, the Digital Single Market, the digitalisation of finance, online platforms, privacy and data protection. ⬈ **[EU 29]**

In their position paper *Automated Decision Making and Artificial Intelligence—A Consumer Perspective* from June 2018 ⬈ **[EU 30]**, BEUC explicitly analyses the increased use of automated decision-making based on algorithms for commercial transactions and its impact on the functionality of consumer markets and societies. It calls for products to be law-compliant by default and that "risks, such as discrimination, loss of privacy and autonomy, lack of transparency, and enforcement failure are avoided."

Looking into the danger of discrimination, it points out that consumers are categorised and profiled with an increasing degree of precision. "The risk of discrimination, intended or unintended, resulting from data input that is not relevant enough to reach a correct conclusion, persists. The user may be deprived of a service or denied access to information, implying severe societal implications." This categorisation leads to the different treatment of each user, either in prices they receive, or deals and services they are offered, based on the association of the customer with a certain group. The report addresses the questions of how ADM processes can be audited, and what appropriate measures for correcting errors could look like.

Commenting on the EU Communication on AI, BEUC criticises that there is no intent to update consumer rights laws with transparency obligations. Such obligations would help to ensure that consumers are informed when using AI-based products and services, particularly about the functioning of the algorithms involved, and their rights to object to automated decisions. On the other hand, BEUC applauds the EU Commission's plan to improve consumer access to their own health data.

/ CLAIRE

CLAIRE (Confederation of laboratories for Artificial Intelligence research in Europe) is an initiative from the European AI community and was publicly launched on June 18, 2018 with a document signed by 600 senior researchers and other stakeholders in Artificial Intelligence. ⬈ **[EU 31]** It seeks to strengthen European excellence in AI research and innovation, and proposes the establishment of a pan-European Confederation of Laboratories for Artificial Intelligence Research in Europe, aiming for a "brand recognition" similar to the European Organisation for Nuclear Research (CERN). Part of CLAIRE's vision is to "focus on trustworthy AI that augments human intelligence rather than replacing it, and that thus benefits the people of Europe."

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ European Association for Artificial Intelligence

The European Association for Artificial Intelligence EurAI (formerly ECCAI) ☑ **[EU 32]** was established in July 1982, as a representative body for the European Artificial Intelligence community. Its aim is to promote the study, research and application of Artificial Intelligence in Europe. The EurAI offers courses, awards and grants for research and dissertations, publishes the journal "AI Communications", and it co-organises the ECAI conference series. In addition, it also participates in the development of recommendations to EU institutions. For example, in January 2018 the European Commission, in collaboration with EurAI, organised a workshop on the European AI landscape. It considered academic, industry, and government AI initiatives, with the aim to share information and strategies for AI across Europe. All countries in this report have member societies in the EurAI.

When asked whether there is currently an ethically correct AI, the president of EurAI, Barry O'Sullivan, had a clear answer: "'No.' […] In principle, the morality of artificial intelligence is a matter for negotiation, for 'there is no consensus on ethical principles—they depend on social norms, which in turn are shaped by cultural values', O'Sullivan added". ☑ **[EU 33]**

/ European Digital Rights (EDRi)

European Digital Rights (EDRi) is an association of and a platform for civil and human rights organisations from across Europe. ☑ **[EU 34]** EDRi's central objective is to promote, protect and uphold civil and human rights in the digital environment. The organisation's goals are to provide policy makers with expert analysis of digital rights issues, foster agenda setting, and coordinate actions between the national and European level in order to ensure that civil society interests and perspectives are included in debates and policy making. EDRi's key priorities for the next years are privacy, surveillance, net neutrality, and copyright reform. The perspective that information technology has a revolutionary impact on our society is the common thread in their campaigns. Dealing with new regulatory measures, EDRi published "A Digestible Guide to Individual's Rights under GDPR".

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

/ euRobotics

euRobotics AISBL (Association Internationale Sans But Lucratif) ☑ **[EU 35]** is a Brussels based non-profit association for stakeholders in European robotics. It was formed to engage from the private side in a contractual public-private partnership, called SPARC, with the European Union as the public side. One of the association's main missions is to collaborate with the European Commission to develop and implement a strategy and a roadmap for research, technological development and innovation in robotics. The "ethical, legal and socio-economic issues in robotics" working group published a first position paper in early 2017. ☑ **[EU 36]**

REGULATORY AND SELF-REGULATORY MEASURES

/ EU General Data Protection Regulation (GDPR) and ADM

"(1) The data subject shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her.
(2) Paragraph 1 shall not apply if the decision: …"—Art. 22 (1), GDPR

It is debated whether the GDPR ⬈ [EU 37] offers the "potential to limit or offer protection against increasingly sophisticated means of processing data, in particular with regard to profiling and automated decision-making". ⬈ [EU 38] But while the GDPR includes a definition of the concept of profiling, mentions ADM explicitly (e.g. in Art. 4, 20, and 22), generally supports the protection of individual interests ("protection of interests of the data subjects", obligations to risk assessment in Art. 35), and widens the rights of "data subjects" impacted by "solely automated" ADM with "legal" or "similarly significant" impact, it remains unclear under what circumstances these protections apply. ⬈ [EU 37]

More precisely, the GDPR defines three criteria as conditions in order for the right of Art. 22 to be applied to ADM: (1) Only when decision-making is fully automated, (2) when it is based on personal data, and (3) when decisions have significant legal consequences or similar effects on the data subject.

It remains a matter of controversy among experts regarding what the GDPR defines as a "decision" or what circumstances and which "legal effects" have to occur for the prohibition to apply. It further does not reflect the diversity of ADM systems already implemented, including various scenarios in which people are involved, who consciously or unconsciously implement ADM or follow the recommendations unquestioningly. ⬈ [EU 39]

Therefore, critics and rights advocates are questioning the scope and effectiveness of the application of Art. 22 to ADM. ⬈ [EU 40] They see little room for manoeuvre when it comes to explicit, well-defined and effectual rights, especially against group-related and societal risks and the impact of automated decision-making systems. Also, concerning its interpretation and application to ADM, it lacks authoritative guidance. ⬈ [EU 39]

EDRi criticises the dilution of the right not to be subjected to automated decisions in Art. 20 of the GDPR leading to a lack of safeguards against the negative effects of profiling on data subjects' privacy, among others. ⬈ [EU 41] Nevertheless, Privacy International points out in their report "Data Is Power: Profiling and Automated Decision-Making in GDPR" that for the first time in EU data protection law the concept of profiling is explicitly defined (Art. 4(4)). It is referred to as "the automated processing of data (personal and not) to derive, infer, predict or evaluate information about an individual (or group), in particular to analyse or predict an individual's identity, their attributes, interests or behaviour". ⬈ [EU 38][7]

Other aspects discussed in regard to the "exceptionally" permissible ADM under the GDPR are transparency and information obligations and the role of data controllers. ⬈ [EU 40]

⬈ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

---

7   EDRi add: "Through profiling, highly sensitive details can be inferred or predicted from seemingly uninteresting data, leading to detailed and comprehensive profiles that may or may not be accurate or fair. Increasingly, profiles are being used to make or inform consequential decisions, from credit scoring, to hiring, policing and national security."

After the approval of the GDPR in 2016, researchers and experts claimed that at least a 'right to explanation' of all decisions made by automated or artificially intelligent algorithmic systems will be legally mandated by the GDPR, while others are very critical about it and see strong limitations. ☑ **[EU 39]** ☑ **[EU 40]**

Taking into account the previously outlined criteria for the GDPR to apply to ADM, critics doubt the legal effect and the feasibility of such a right because of a lack of precise language and of explicit and well-defined rights and safeguards against automated decision-making. ☑ **[EU 39]** Moreover, these obligations—including systemic and procedural provisions and regulatory tools given to data controllers and data protection authorities, e.g. to carry out impact assessments and data protection audits—focus on the protection of individual rights and freedoms.[8] The GDPR transparency rules do not include mechanisms for an external deep look into the ADM systems necessary to protect group-related and societal interests such as non-discrimination, participation or pluralism. ☑ **[EU 39]**

This means that there is a high probability that the GDPR's provisions specific to ADM only apply in the rarest of cases; systems preparing human decisions and giving recommendations may still be used. But there is a lively debate about what the impact of the regulation on the development of ADM will look like. Complementary approaches to strengthen measures within the GDPR and alternative regulatory tools are discussed to rectify already implemented ADM systems, e.g. by using regulation already in place (consumer protection law, competition law, and media law). ☑ **[EU 40]**

The European Data Protection Board (EDPB) ☑ **[EU 42]**, which replaced the Article 29 Working Party (WP29) on the Protection of Individuals with regard to the Processing of Personal Data, is responsible for the development and publication of 'guidelines, recommendations and best practices' on the GDPR from the EU side. ☑ **[EU 42]**

The European data protection bodies'[9] Guidelines on Automated individual decision-making and Profiling for the purpose of Regulation 2016/679 (developed by WP29, endorsed by the EDPB, last revised and adopted in February 2018) suggest not only more comprehensive definitions of both concepts and their possible implications, but also general and specific provisions. It provides examples for conditions necessary for the protection to apply. According to those guidelines, a "legal or similarly significant impact" is manifest in the cancellation of a contract, denial of social benefits, access to education, eligibility for credit, or if it leads to the exclusion or the discrimination of individuals and affects the choices available to the subject. ☑ **[EU 43]**

/ Police Directive

The EU data protection reform package of 2016, which included the GDPR, also involved a directive on data protection in the area of police and justice ☑ **[EU 44]** as a *lex specialis*, adopted on May 5, 2016, applicable as of May 6, 2018. Where, however, the GDPR is

---

8   Again, the implicit "right to explanation of the system functionality", or "right to be informed" is restricted by the interests of data controllers and the interpretation and definition of ADM in Art. 22.

9   The Article 29 Data Protection Working Party (WP29) was set up in 1996 by the Directive 95/46/EC as an independent European advisory body on data protection and privacy. With the GDPR in place, the European Data Protection Board (EDPB) replaced the WP29, endorsing its guidance provided on the GDPR. The EDPB has a role to play in providing opinions on the draft decisions of the supervisory authorities. For this purpose, it can examine—on its own initiative or on the request of one of its members or the European Commission—any question covering the application of the GDPR.

directly applicable as regulation, the directive[10] had first to be transposed into national law, allowing more space for variations at the national level. It is also intended to facilitate cross-border cooperation and regulate the exchange of data, e.g. with Interpol.

The directive regulates the processing of personal data by criminal law enforcement authorities and further agencies responsible for the prevention, investigation, detection and prosecution of criminal offences. It is intended to ensure, in particular, that the personal data of victims, witnesses, and suspects of crime are duly protected.

The directive applies key principles and provisions of the GDPR to European criminal law enforcement authorities, though with adjusted requirements and specified exceptions. Law enforcement data protection officers, for example, are to be designated and must face the same responsibilities as others with this position under the GDPR. On the other hand, a more complex requirement is the differentiation of personal data based on facts from those based on assessments (Art. 7).

Regarding automated individual decision-making, it states in Art. 11:

1. "Member States shall provide for a decision based solely on automated processing, including profiling, which produces an adverse legal effect concerning the data subject or significantly affects him or her, to be prohibited unless authorised by Union or Member State law to which the controller is subject and which provides appropriate safeguards for the rights and freedoms of the data subject, at least the right to obtain human intervention on the part of the controller.
2. Decisions referred to in paragraph 1 of this Article shall not be based on special categories of personal data referred to in Article 10, unless suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place.
3. Profiling that results in discrimination against natural persons on the basis of special categories of personal data referred to in Article 10 shall be prohibited, in accordance with Union law." ☑ **[EU 44]**

One concern is the applicability of the Police Directive to public–private partnerships. Taking into account the complexity and legitimacy of these forms of cooperation, e.g. when outsourcing the technological implementation of measures for combatting cybercrime and law enforcement data processing, it is not clear if these structures are subject to the data protection-related regimes of the GDPR and the Police Directive. ☑ **[EU 45]**

/ Strengthening trust and security

The European Commission's initiatives to improve online security, trust and inclusion moreover comprise (1) the ePrivacy Regulation ☑ **[EU 46]** ☑ **[EU 47]** (2) the Cybersecurity Act[11] ☑ **[EU 48]** and (3) the Safer Internet Programme. ☑ **[EU 49]**

☑ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

---

10   The directive repeals the Council Framework Decision 977/2008/JHA.

11   So far, it is planned to establish a „European Cybersecurity Certification Group" consisting of representatives of the national authorities responsible for cyber security certification.

/ EU Data Protection Bodies

European Data Protection Supervisor

In addition to the European Data Protection Board (see above, under GDPR), the European data protection bodies include the European Data Protection Supervisor (EDPS). The EDPS, as an independent supervisory authority, has the responsibility to monitor the processing of personal data by EU institutions and bodies, advise on policies and legislation that affect privacy, and cooperate with similar authorities either at the national or international level to ensure consistent data. ☑ **[EU 50]**

The Supervisor and Assistant Supervisor were appointed in December 2014 with the specific remit of being more constructive and proactive. In March 2015, they published a five-year strategy ☑ **[EU 51]** setting out how they intended to implement this remit, and to be accountable for doing so. In the strategy they first recognise:

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

> "Big data challenges regulators and independent authorities to ensure that our principles on profiling, identifiability, data quality, purpose limitation, and data minimisation and retention periods are effectively applied in practice.
>
> Big data that deals with large volumes of personal information implies greater accountability towards the individuals whose data are being processed. People want to understand how algorithms can create correlations and assumptions about them, and how their combined personal information can turn into intrusive predictions about their behaviour." ☑ **[EU 51]**

The EDPS' action plan to tackle these and more challenges include the following:

- Promoting technologies to enhance privacy and data protection
- identifying cross-disciplinary policy solutions
- increasing transparency, user control and accountability in big data processing
- developing an ethical dimension to data protection
- mainstreaming data protection into international agreements
- speaking with a single EU voice in the international arena
- adopting and implementing up-to-date data protection rules
- increasing the accountability of EU bodies processing personal information, and
- facilitating responsible and informed policymaking

Their objective and mandate is to ensure that data protection is integrated into proposals for legislation that affects privacy and personal data protection in the EU. The European Commission consults the EDPS before it adopts a proposal for new legislation that is likely to have an impact on individuals' right to the protection of their personal data. The EDPS also provides advice on EU initiatives that are not legally binding so-called EU soft law instruments. It issues comments and opinions related to proposals for legislation that are addressed to all three EU institutions involved in the legislative process. In addition, it publishes recommendations or comments on their own initiative, when appropriate, and when there is a matter of particular significance. ☑ **[EU 52]**

The EDPS may also intervene before the EU courts either at the Court's invitation or on behalf of one of the parties in a case to offer data protection expertise. Moreover, the EDPS monitors new technologies or other societal changes that may have an impact on data protection.

On July 10, 2018, the EDPS issued an Opinion on the European Commission's Proposal for a new directive on the re-use of Public Sector Information (PSI). It provides specific recommendations on how to clarify the relation and coherence of the PSI Directive with the exceptions outlined in the GDPR. [EU 53]

In its opinion on the proposal, the EDPS points out the relevance of the revision of the PSI Directive as part of the EU vision on "Good Big Data" and emphasises how the smart use of data, including its processing via Artificial Intelligence, can have a transformative effect on all sectors of the economy. At the same time, the EDPS demands that the legislator better address stakeholder concerns related to the necessary protection of personal data, especially in sensitive sectors such as health, when they take the decision on the re-use of PSI (for example, clarifying the risks of re-identification of anonymised data and the safeguards against those risks).

In that context, the EDPS recalls the data protection-relevance of the key principles that, according to the European Commission, should be respected in the context of data re-use, namely (i) minimised data lock-in and ensureing undistorted competition; (ii) transparency and societal participation on the purpose of the reuse vis-à-vis the citizens/data subjects as well as transparency and clear purpose definition between the licensor and the licensees; (iii) data protection impact assessments and appropriate data protection safeguards for reuse (according to a 'do no harm'—under the data protection viewpoint—principle).

Additionally, it provides for further recommendations on anonymisation and its relation to costs and data protection. It also focuses on a data protection impact assessment (DPIAs) for sensitive sectors, such as healthcare, while taking into account an 'acceptable re-use policy'.

Some of the EDPS's output relevant to ADM includes:

- As part of its Ethics Initiative, the EDPS conducted a public consultation, lasting from June to July 2018, showing the need to re-think the role of data in the digital era along with questions like "What does the right to privacy mean in an age of continuous and ubiquitous tracking, measuring, and profiling? What does data protection mean in the age of big data processing and its apparent and real opportunities? How can human dignity and autonomy be held? And how can the benefits brought about by new digital technologies be equitably shared among all?" [EU 54]
- The Declaration on Ethics and Data Protection in Artificial Intelligence, adopted at the 2018 International Conference of Data Protection and Privacy Commissioners, endorses guiding principles in regard to AI—including elaborations on fairness, accountability, systems transparency and intelligibility, ethics and privacy by design, empowerment and public engagement and the reduction and mitigation of unlawful bias and discrimination [EU 55]
- An early opinion by the EDPS from 2014 analyses the EU Commission's proposal for a Regulation of the European Parliament and of the Council on a European network of Employment Services, workers' access to mobility services and the further integration of labour markets, among others, on the use of ADM in job matching at the EURES portal [EU 56]

/ EU employment equality law

On the basis of three directives (2000/43/EC, 2000/78/EC, 2002/73/EC) and various court decisions especially dealing with Art. 157 of the Treaty on the Functioning of the European Union (TFEU), a general framework for equal treatment in employment and occupation has been established in the EU. This framework could be applied to protect workers in a situation where work-related decisions are not taken by a human being, but (semi-)automatically by an algorithm, since this is a potential source of discrimination. The framework doesn't just include formally employed persons, as many national regulations do, but has a wider approach, so it may also be applicable to platform workers as well. In addition, the recent Egenberger decision of the ECJ established that anti-discrimination applies to all aspects laid down in Article 21 of the Charter of Fundamental Rights of the European Union, namely sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age, or sexual orientation. Since this list is not conclusive ("Any discrimination based on any ground such as ..."), there is room for other grounds of discrimination that can be adressed.

Related documents:

- Charter of Fundamental Rights of the European Union (2007/C 303/01 ☐ **[EU 57]**)
- Council Directive 2000/43/EC of 29 June 2000 implementing the principle of equal treatment between persons irrespective of racial or ethnic origin ☐ **[EU 58]**
- Council Directive 2000/78/EC of 27 November 2000 establishing a general framework for equal treatment in employment and occupation ☐ **[EU 59]**
- Directive 2006/54/EC of the European Parliament and of the Council of 5 July 2006, repealing Directive 2002/73/EC on the implementation of the principle of equal opportunities and equal treatment of men and women in matters of employment and occupation (recast) ☐ **[EU 60]**
- Platform Work, Algorithmic Decision-Making, and EU Gender Equality Law ☐ **[EU 61]**
- ECJ Case C-414/16, Vera Egenberger v Evangelisches Werk für Diakonie und Entwicklung e.V. ☐ **[EU 62]**

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ Financial market directive – MiFID 2 and MiFIR

In June 2014, the European Commission adopted new rules revising the Markets in Financial Instruments Directive (2004/39/EC). The MiFID framework has been applicable since January 3, 2018. These new rules consist of a directive (MiFID 2) and a regulation (MiFIR). ☐ **[EU 63]** MiFID 2 aims to reinforce the rules on securities markets by, among other things, introducing rules governing high frequency algorithmic trading (HFAT).[12]

The new legislative framework is supposed to "ensure fairer, safer and more efficient markets and facilitate greater transparency for all participants" ☐ **[EU 65]**. The protection of investors is strengthened through the introduction of new (reporting) requirements, tests and product governance, and independent investment advice, as well as the improvement of requirements in several areas. These include the responsibility of management bodies, inducements, information and reporting to clients, cross-selling, remuneration of staff, and best execution.

---

12 "High frequency algorithmic trading (HFAT) is a subset of algorithmic trading. Algorithmic trading uses computer algorithms to automatically determine parameters of orders such as whether to initiate the order, the timing, price or how to manage the order after submission, with limited or no human intervention. The concept does not include any system used only for order routing to trading venues, processing orders where no determination of any trading parameters is involved, confirming orders or post-trade processing of transactions." ☐ **[EU 64]**

HFAT investment firms and trading venues are facing a set of organisational requirements, e.g. to store time-sequenced records of their algorithmic trading systems and trading algorithms for at least five years. To enable monitoring by Member State competent authorities, the European Securities and Markets Authority (ESMA) proposes that the records should "include information such as details of the person in charge of each algorithm, a description of the nature of each decision or execution algorithm and the key compliance and risk controls." ⧉ **[EU 64]**

Further provisions regulating the non-discriminatory access to trading venues, central counterparties (CCPs) and benchmarks are designed to increase competition.

/ European Standardisation Organisations (CEN, CENELEC, ETSI)

The three European Standardisation Organisations[13], CEN ⧉ **[EU 67]**, CENELEC ⧉ **[EU 68]** and ETSI ⧉ **[EU 69]** are officially recognised as competent in the area of voluntary technical standardisation. CEN, CENELEC and ETSI enable businesses to comply with relevant directives and regulations through the development of Harmonised Standards (HS) which are published in the Official Journal of the European Union (OJEU).

As contributors to the development of the EU Digital Single Market, the European Standardisation Organisations are tackling fields of work such as additive manufacturing, blockchain, artificial intelligence and cybersecurity.

⧉ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

/ Connected and Automated Mobility – new rules on autonomous driving

Following the previous 'Europe on the Move' initiative of May 2017, the European Commission in 2018 announced a strategy for automated and connected transportation in Europe, the so-called 3rd Mobility Package. ⧉ **[EU 70]**

The EU is planning ⧉ **[EU 71]** to adopt a new policy recommendation by the end of 2018 / beginning of 2019, setting out the legal framework for the communication between autonomous vehicles and road infrastructure. This is to be achieved by means of so-called "Cooperative Intelligent Transport Systems" (C-ITS). These C-ITS would be installed as boxes in vehicles, cars and roads, and connected to traffic control centres. In particular, the EU turned its attention to Japan, where C-ITS has been operating successfully since 2016.

The regulation is supposed to ensure that drivers of vehicles equipped with this technology will be informed of "13 dangerous situations in stationary and moving traffic". Vehicle movements are to be coordinated in order to, for example, trigger braking manoeuvres and "drastically reduce" frequently occurring accident sequences. In the second phase, further infrastructure is to be integrated into 'intelligent' data exchange, like charging stations, parking spaces and park and ride areas.

By mid-2019, a "100% guaranteed reliability of the warning notices" is intended to create more legal certainty for automobile manufacturers, road operators and companies in the

---

13   The European Union (EU) Regulation (1025/2012) that settles the legal framework for standardisation, has been adopted by the European Parliament and by the Council of the EU, and entered into force on 1 January 2013. ⧉ **[EU 66]**

telecommunications sector. The standards developed include specifications and safe-guards for data protection and cyber security.

After about four weeks of consultation with Member States and manufacturers specialising in autonomous driving ⬈ **[EU 72]** ⬈ **[EU 73]**, the Council and the European Parliament still have to approve the regulation, which is expected to be operational by mid 2019.

ADM IN ACTION

/ DANTE anti-terrorism project

DANTE ("Detecting and analysing terrorist-related online contents and financing activities") is an experimental project, funded by the European Commission within the Horizon2020 programme, and aimed at using automated decision-making against terror-ism. ⬈ **[EU 74]** Eighteen EU countries are involved. DANTE is described as a "framework" that supplies "innovative knowledge mining, information fusion, and automated reasoning techniques and services" for the discovery and analysis of terrorist networks by law en-forcement and intelligence agencies. It includes "automated functionalities" as wide-rang-ing as: "detection and monitoring of sources of relevant terrorist-related data in surface/deep Web, and dark nets; accurate and fast detection, analysis, categorisation of suspect terrorist related multi-language contents; large-scale temporal analysis of terrorism trends; real-time summarisation of multilingual and multimedia terrorist-related content; detection of disinformation in online content; detection and monitoring of relevant indi-viduals and linking pseudonyms with the original authors; accurate and fast identification of terrorist online communities and groups; capturing, storing and preserving relevant data for further forensic analysis". Results are yet to be published.

In the meantime, it was announced that DANTE is collaborating with another Horizon 2020 project, TENSOR. ⬈ **[EU 75]** Together, they aim to develop a platform offering Law Enforce-ment Agencies planning and prevention applications for early detection of terrorist activi-ties, radicalisation and recruitment by the means outlined above. TENSOR is said to have "a work stream dedicated to the ethical, legal and societal impact".

/ EU Border Regime & interoperability of EU information systems

In recent years, the European Union has been proposing and adopting mechanisms and initiatives to establish an "integrated smart border management" system. ⬈ **[EU 76]** At the same time, it has launched a process towards the interoperability of (existing and future) large-scale EU information systems. ⬈ **[EU 77]** This is aimed at integrating instruments for data processing and decision-making systems in the fields of asylum and immigration, border management, and law enforcement cooperation. These developments represent the gradual moving away from a "country-centric" approach towards a "person-centric" approach . ⬈ **[EU 78]** Though strongly criticised by civil society and data protection bodies (see below in this paragraph), and accompanied by the request for technological reviews ⬈ **[EU 79]**, the implementation of an overarching interoperable smart border management system is on its way.

eu-LISA, the "European Agency for the Operational Management of large-scale IT Sys-tems in the Area of Freedom, Security and Justice", is now managing the "strengthened" databases and applications VIS, SISII and EURODAC together. ⬈ **[EU 80]** This is leading to

the creation of a "biometric core data system" ⤢ **[EU 81]**, with three more systems under construction or currently being discussed:

- A new centralised version of the European Criminal Records Information System which will also include third-country nationals (ECRIS-TCN).
- The adapted European Travel Information and Authorisation System (ETIAS). ⤢ **[EU 82]**. This is a pre-travel authorisation system that comes into force in 2021. It includes an "individual risk assessment" process based on a "background and eligibility questionnaire" for visa-exempt travellers. It cross-checks with EU databases and the "ETIAS watchlist". The applications are processed automatically.[14]
- Complemented by the Entry-Exit System (EES) ⤢ **[EU 83]**, which is currently under development and will be operational by 2020, EES will be used by border guards and consular officers. Member States law enforcement authorities and Europol will also have access to it.

eu-LISA is the agency mandated to implement "technical upgrades" of these IT systems. It runs pilot projects and training, e.g. on the use of identification and verification technologies for border control.

Reflecting on the interoperability of EU information systems to freedom, security and justice the European Data Protection Supervisor stresses "that interoperability is not primarily a technical choice, it is first and foremost a political choice to be made, with significant legal and societal implications in the years to come". It sees a "clear trend to mix distinct EU law and policy objectives". ⤢ **[EU 84]** It follows the criticism of MEP Marie-Christine Vergiat who claimed that a "presumption of irregularity" underlining this system replaces the assumption of innocence. ⤢ **[EU 85]** Michael O'Flaherty, the director of the European Union Agency for Fundamental Rights, addressed the High-Level Expert Group on Information Systems and Interoperability and scrutinised the effect of "flagged" hits—this is the only knowledge that an entry for a specific person exists in a specific database, for example in the European Criminal Records Information System (ECRIS-TCN)—and if a "flagged" hit may further influence decisions taken about an individual. O'Flaherty also underlines that "there is the risk of discriminatory profiling. The data contained in IT systems can be used for risk assessment or profiling. Risk assessment or profiling is not a violation of fundamental rights, but discriminatory profiling is. The chance of this taking place increases if IT systems are interoperable, as several data categories revealing sensitive data such as, race, ethnicity, health, sexual orientation, and religious beliefs can then be accessed simultaneously for profiling purposes." ⤢ **[EU 86]**

With this in mind, see the paragraph on iBorderCtrl, a project on an "Intelligent Portable Border Control System" below.

---

14   When there is a hit in one of the EU security databases or a question is answered positively, the data will be manually checked and the risks individually assessed within four weeks. Persons posing a risk will be refused entry, having the right to appeal.

/ iBorderCtrl – "Intelligent Portable Border Control System"

iBorderCtrl ⬀ **[EU 87]**[15] is a system designed to screen non-EU nationals at EU borders. It is supposed to determine whether a person is telling the truth in an automated interview with a virtual border guard, conducted before the person arrives at the border, i.e. from home. If they pass this screening process, they can pass the border without further investigation. If there is suspicion that a person is lying, then biometric information is taken at the border control point, such as fingerprints and palm vein images, and the person is passed on to a human agent who will review the information and make an assessment. iBorderCtrl is currently being tested in Hungary, Greece, and Latvia, where those countries have borders with non-EU countries. ⬀ **[EU 88]**

⬀ **LINKS:** *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

In the first phase—pre-screening with automated checks—third-country nationals register online, upload pictures of their passport, visa, and proof of funds, then they use a webcam to answer questions from a computer-animated border guard avatar. This is described as "a short, automated, non-invasive interview, subject to lie detection" and aimed to "improve performance in comparison to human agents, as it correctly adapts to traveller's profiles"— meaning it is personalised to the traveller's gender, ethnicity and language. In a "unique approach to 'deception detection'" the system "analyses the micro-gestures of travellers to figure out if the interviewee is lying." ⬀ **[EU 89]**

Besides the fundamental question of the scientific accuracy of lie detectors, the group of 32 people who tested the avatar was not representative and contained the risk of racial bias: the majority were men (22), and they were mainly white Europeans. Only 10 had an Asian or Arab background. After asking 13 questions, the average accuracy of the system in relation to a single question is 75%, meaning there is a potential 25% error rate. ⬀ **[EU 90]**

Travellers at this stage are further informed "of their rights and travel procedures" and provided with "advice and alerts to discourage illegal activity".

At stage two—at the border itself—border officials "automatically cross-check information, comparing the facial images captured during the pre-screening stage to passports and photos taken on previous border crossings." Afterwards, the risk score of the traveller is supposed to be recalculated. "Only then does a border guard take over from the automated system." ⬀ **[EU 91]**

One of the stated aims of the system is "to reduce the subjective control and workload of human agents and to increase the objective control with automated means that are non-invasive and do not add to the time the traveller has to spend at the border". The application further foresees the creation of an additional layer for "bona fide travellers, especially regular travellers into a Schengen-wide frequent traveller programme including a reward system based on number of successful crossings and trouble-free stay". ⬀ **[EU 87]**

The stated technological framework of the iBorderCtrl system ⬀ **[EU 92]** is to "empower —through technology—the border guard" and it involves:

---

15   The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 700626. "The piloting deployment will end in August 2019 after approximately nine months duration. As the hardware and software systems under development in the iBorderCtrl project are not yet authorised law enforcement systems, national legislative authorisation of data processes are not applicable. Informed consent is required to involve test participants and to process their data." ⬀ **[EU 87]**

Kristina Penner

is the Executive Advisor at AlgorithmWatch. Her research interests include ADM in social welfare systems, social scoring and the societal impacts of ADM as well as the development of participatory and empowering concepts. Her analysis of the EU border management system builds on her previous experience in research and counselling on the implementation of European and German asylum law. Further experience includes projects on the use of media in civil society and conflict sensitive journalism as well as stakeholder involvement in peace processes in the Philippines. She holds a master's degree in International Studies / Peace and Conflict Research from Goethe University in Frankfurt.

- **An Automatic Deception Detection System (ADDS)**[16] that performs, controls and assesses the pre-registration interview that is personalised to suit the gender and language of the traveller. ADDS quantifies the probability of deceit in interviews by analysing interviewees' non-verbal micro expressions.
- **A Biometrics Module** for the biometric identity validation, comparing data stored in databases (legacy systems in the case of fingerprints and creation of a baseline database for palm vein images).
- **A Face Matching Tool (FMT),** including video and photo to create a biometric signature in order to provide a matching score.
- **A Document Authenticity Analytics Tool (DAAT)** for the verification procedure of travel documents, which are examined by DAAT against fraud characteristics in an automated way. A matching score concerning the authenticity of documents is then derived.
- **An External Legacy and Social interfaces system (ELSI),** crosschecking the traveller's information from social media or legacy systems, such as SIS II.
- **A Risk Based Assessment Tool (RBAT),** utilising risk based approaches to intelligently aggregate and correlate all the data collected and the estimated risk. It then classifies travellers to support the decision of the border guard. This includes a systematic process to stimulate compliance by compressing all the data into meaningful actionable risk scores.
- **An Integrated Border Control Analytics Tool (BCAT)** for advanced post-hoc analytics.
- **A Hidden Human Detection Tool (HHD)** to detect people inside various vehicles.

iBorderCtrl states that "regarding the expected accuracy it would be wrong to expect 100% accuracy from any AI-based deception detection technology, no matter how mature", iBorderCtrl therefore relies "on many components that address various aspects of the border control procedures, and each provides its own risk estimation for the traveller". The system then "synthesises a single risk score from a weighted combination of components". Emphasising the "human-in-the-loop principle", the makers conclude that "it is highly unlikely that an overall system of which ADDS is a part will lead to 'an implementation of a pseudoscientific border control.'" ⧉ **[EU 92]**

According to the scientists involved, EU funding ensures that economic interests do not play a role. Nevertheless, the so-called "success story" of the Commission on the project ends with:

> "[...] 'in light of the alarming terror threats and increasing terror attacks taking place on European Union soil, and the migration crisis' [...], the partner organisations of IBORDERCTRL are likely to benefit from this growing European security market—a sector predicted to be worth USD 146 billion (EUR 128 bn) in Europe by 2020." ⧉ **[EU 91]**

---

16   On the use of ADDS, apart from the research project, the website adds that "in a real border check [it] can not be based on informed consent. A legal basis would be needed, which at present does not exist in the applicable European legal framework."

In 2016, a local police depart-
ment on the Belgian coast
started implementing a predic-
tive policing system. According
to the police the crimes the
system is most effective at
predicting are burglaries and
vehicle theft.

The Belgian Liga voor
Mensenrechten (Human
Rights League) works on pri-
vacy and other human rights
issues and is increasingly
looking into new technolo-
gies, especially in the context
of its Big Brother Awards.

WESTKUST/KOKSIJDE,
DE PANNE,
NIEUWPOORT

GENT

BRUSSELS

LEUVEN

The Belgian Chamber of
Representatives in 2018
adopted a resolution
to preventatively ban
fully automated weapons
('killer robots').

The public employment
service of Flanders developed
algorithms to analyse online
behaviour of job seekers, to
allow early and "more efficient"
intervention.

BELGIUM

BY ROSAMUNDE VAN BRAKEL

Belgium has several levels of government and debates are spread out over the different governments: federal, regional (Flemish, Walloon and Brussels-Capital region government) and the community governments (Flemish, French and German). In general, the current Belgian political discourse uses the terms 'Big Data' and 'Artificial Intelligence' (AI). In general, the governmental strategies concerning digitisation and AI are embedded in an economic discourse and are about increasing jobs and supporting companies in the context of Industry 4.0. Although automation and digitisation has been going on for a long time, in Belgium this has always been characterised by problems with implementation. The result is that Belgium is rather behind the rest of Europe, especially when it comes to digitisation of the public sector.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/automating-society

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ Digital Belgium: Digital Agenda – Federal Government

The federal government does not have a specific strategy concerning AI or ADM, but they did launch "Digital Belgium: Digital Agenda" in 2015. According to this agenda, by 2020 it should be possible for Belgium to get into the top three of the European *Digital Economy and Society Index*, for 1,000 new start-ups to take root in Belgium, and for the digital revolution to deliver 50,000 new jobs in a variety of sectors.

"Digital Belgium" is a plan that outlines the long-term digital vision for the country and translates this into five priorities of the federal government:

- Digital infrastructure
- Digital confidence and digital security
- Digital government
- Digital economy
- Digital skills and jobs ⬈ **[BE 1]**

Recently, Belgian prime minister Charles Michel commented on AI, saying that he "is convinced that we will need new professions in the future and that they will be made possible by this technological evolution", and that he is "convinced that artificial intelligence is an opportunity for quality of life, to advance the quality of medicine, telecommunications, and to raise the standard of living on this earth." ⬈ **[BE 2]**

The Belgian Privacy Commission published 33 recommendations about Big Data in 2017. Most of the recommendations refer to the GDPR, especially when it comes to automated or semi-automated decision-making. ⬈ **[BE 3]**

/ Radically digital Flanders – Flemish Government (Region and Community)

The Flemish government launched its digital strategy: Vlaanderen radicaal digitaal (Radically digital Flanders). It is inspired by the Digital Agenda of the Federal government and has the same priorities. ⤢ **[BE 4]**

The Flemish government also has a specific programme on Artificial Intelligence that looks at how AI can improve government services. Five programme directions are identified as:

- The human computer: **digital = super handy,** 'thinks and integrates' like people thanks to chatbots and conversational platforms
- The computer assistant: **hyperpersonalisation in government services** by collecting data about citizens—and starting from that knowledge—offer a more personal government experience.
- The super-quick (proactive) computer: **a quick smart government as a result of text, language and image recognition** and other pattern recognitions.
- The autonomous computer: **more with less.** Through automation of tasks, civil servants can do other work.
- The moral computer: **digital ethics**. We are actively following European initiatives. ⤢ **[BE 5]**

The Flemish Minister of Innovation Philippe Muyters announced that the government will invest €30 million in AI. In his "AI action plan", he presents three main goals: fundamental research, applications for industry, and framing policy concerning education, sensitisation and ethics. PwC, an international consultancy, has been appointed to do an international benchmarking exercise of Flanders to get a better picture of where Flanders is at concerning AI. According to Muyters, the "potential societal and economic impact of AI is enormous. For Flanders, the biggest opportunities lie in the first instance in personalised healthcare, smart mobility and industry 4.0. If we tackle this evolution quickly and smartly, we can make sure Flanders will reap all the benefits." ⤢ **[BE 6]**

The Flemish government states that research programmes that are proven to be of international excellence will be strengthened and deepened, that the government will make "clear choices on the basis of excellence, so that budgets will not be fragmented but instead will be invested in areas where there is the highest potential. Special attention will go to leading AI-technology platforms with clear market potential."

The Innovation Ministry believes that Flanders "can be one of the frontrunners for the application of AI in the business community. This can be done, not by inventing everything, but rather by functioning as a living lab for Flemish and international applications." So-called priority clusters and Vlaio (the Flemish Bureau for Innovation and Entrepreneurship) are supposed to "take care of knowledge sharing and to establish a network to follow AI trends and translate these to Flemish companies."

The Flanders government states that there is a need for a broad sensitisation to the disruptive potential of AI technology: "Both in education and in the corporate world people are working at installing permanent training provisions. In addition, an AI think tank will be established to examine the ethical implications that AI entails."

Several events have taken place in 2018 under the direction of the minister. In July a "stakeholders forum on Artificial Intelligence" was organised. ⤢ **[BE 7]** In September, a conference and exhibition took place to show the potential of AI: SuperNova ⤢ **[BE 8].**

⤢ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/automating-society*

A parliamentary question in the Flemish Parliament on October 3, 2018 discussed the above-mentioned plan of minister Muyters. High on the Flemish political agenda is the need to develop new degrees at universities and training in Artificial Intelligence ☑ **[BE 9].** A new masters degree in Artificial Intelligence has already been launched at the Katholieke Universiteit Leuven ☑ **[BE 10].**

/ Digital Plan – Walloon Government

In December 2015, the Walloon government adopted its digital *Plan du Numérique* ("digital plan") ☑ **[BE 11]**. Its main goal is to become a major Industry 4.0 player and a forerunner in the digital revolution. It is inspired by the Digital Agenda of the Federal Government and has the same priorities.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ Ligue des droit de l'homme

The Ligue des droit de l'homme is a Walloon non-profit organisation for human rights in Belgium. They have a commission looking into the consequences of new technology for human rights. The commission initiates actions and activities that allow it to get in touch with and to react to the population and / or to motivate the creation of citizen initiatives. To this end, the commission is responsible, alone or in collaboration with other actors, for setting up activities or projects. The commission is also responsible for examining files, drafting working papers, articles and position papers, setting up or intervening in conferences or other awareness-raising activities, initiating or participating in action plans, bringing challenges to the courts and confronting the public authorities on the themes within its competence. ☑ **[BE 12]**

/ Liga voor Mensenrechten

The Liga voor Mensenrechten protects human rights by denouncing structural and incidental violations to create a societal foundation for human rights in Belgium. They do this by informing, taking action, and going to court. They work on privacy and other human rights issues related to new technologies, and they organise the Big Brother Awards. They do not work specifically on ADM or Artificial Intelligence, but they are increasingly looking at these technologies especially in the context of the Big Brother Awards. ☑ **[BE 13]**

/ Privacy Salon

Privacy Salon is a non-profit organisation, which aims at sensitising and critically informing the broader public, policy makers and industry in Belgium, Europe, and beyond about privacy, data protection and other social and ethical issues that are raised with the introduction of new technologies in society. Privacy Salon organises the annual CPDP conference where several panels focus on ADM. One of the main themes that is being worked on by the organisation is algorithmic discrimination and algorithmic decision-making. More specifically they are organising events on this theme including an art exhibition and a workshop on Algorithms and Society. ☑ **[BE 14]**

/ Royal Flemish Academy of Belgium for Science and the Arts (KVAB)

The KVAB published an opinion piece in 2017 on Artificial Intelligence. The main purpose of this document was "to inform the public as objectively as possible and to propose a series of conclusions and recommendations to concerned parties in order to deal with AI and ensure that our community can properly benefit from its huge opportunities, as well as get an insight into the risks and what to do about them." Also, The Class of Natural Sciences (KNW) of the KVAB has started a working group to study the impact of AI in Flanders. ⤴ **[BE 15]**

REGULATORY AND SELF-REGULATORY MEASURES

/ Belgian Law concerning the protection of data of natural persons in relation to the processing of personal data

The most important law regulating automated decision-making is the data protection regulation. The Belgian version of the GDPR, and the replacement of the 1992 Belgian Privacy Law ⤴ **[BE 16]**, came into force on September 5, 2018. The law applies to every fully or partially automated processing of personal data, and also to the processing of personal data which is not automated, but which is included in a file or will be included in a file. ⤴ **[BE 17]**

/ 'Killer robots'

The Belgian Chamber of Representatives adopted a resolution in 2018 to have a preventative ban on fully automated weapons ('killer robots'). ⤴ **[BE 18]**

ADM IN ACTION

/ Algorithmic work activation

The public employment service of Flanders, VDAB, together with the Katholieke Universiteit Leuven and Vlerick Business school developed algorithms that provide insight into the way people search for jobs on their website. ⤴ **[BE 19]** The system analyses thousands of job seeker files and looks at the click behaviour of people who are looking for jobs on the VDAB website. According to the VDAB, this process has an important predictive value concerning long-term unemployment. The information is supposed to allow for early and more efficient intervention. One goal for the VDAB is to see if click behaviour analysis can be used to control the active search behaviour of the job seeker. The job seeker who is not active enough online would then be invited for an interview and the next step would be a penalty ⤴ **[BE 20].** Another application would be similar to Amazon's recommendation system. On the basis of the huge amounts of data VDAB collects, it could then provide the person with a list of recommended jobs and present potential employers to the right candidates. According to the VDAB, by using these data-driven methods it is possible to improve personal guidance of jobseekers. ⤴ **[BE 21]**

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ Predictive policing

In 2016, a local police zone on the Belgian coast started implementing predictive policing software. The chief commissioner claims that since the start of the project, criminality has gone down by 40 %. According to the police, the types of criminality that the predictions are the most effective at are burglaries and vehicle theft as there is a lot of data available about these crimes that can be analysed by the software. Via the data that the police receives, they claim that they can predict in which neighbourhoods it is more likely that a burglary will take place. On the basis of this prediction they will send out an intervention team. The chief commissioner wants to expand the system by interconnecting the software with Automatic Number Plate Recognition (ANPR) cameras. ⧉ **[BE 22]** ⧉ **[BE 23]**

In 2016 the Belgian federal government invested in the iPolice system to centralise all police data in the cloud. The system should be operational by 2020. This is a cooperation between the Ministry of Home Affairs, the Digital Agenda and the Ministry of Justice. In an answer to a parliamentary question, the minister of Home Affairs, Jambon, stated on October 27, 2016: "The new technologies should make possible a better linkage, sharing and analysis of information in a quick way. The police should work and act on the basis of an integral analysis of structured and unstructured data, from internal and external available data." ⧉ **[BE 24]**
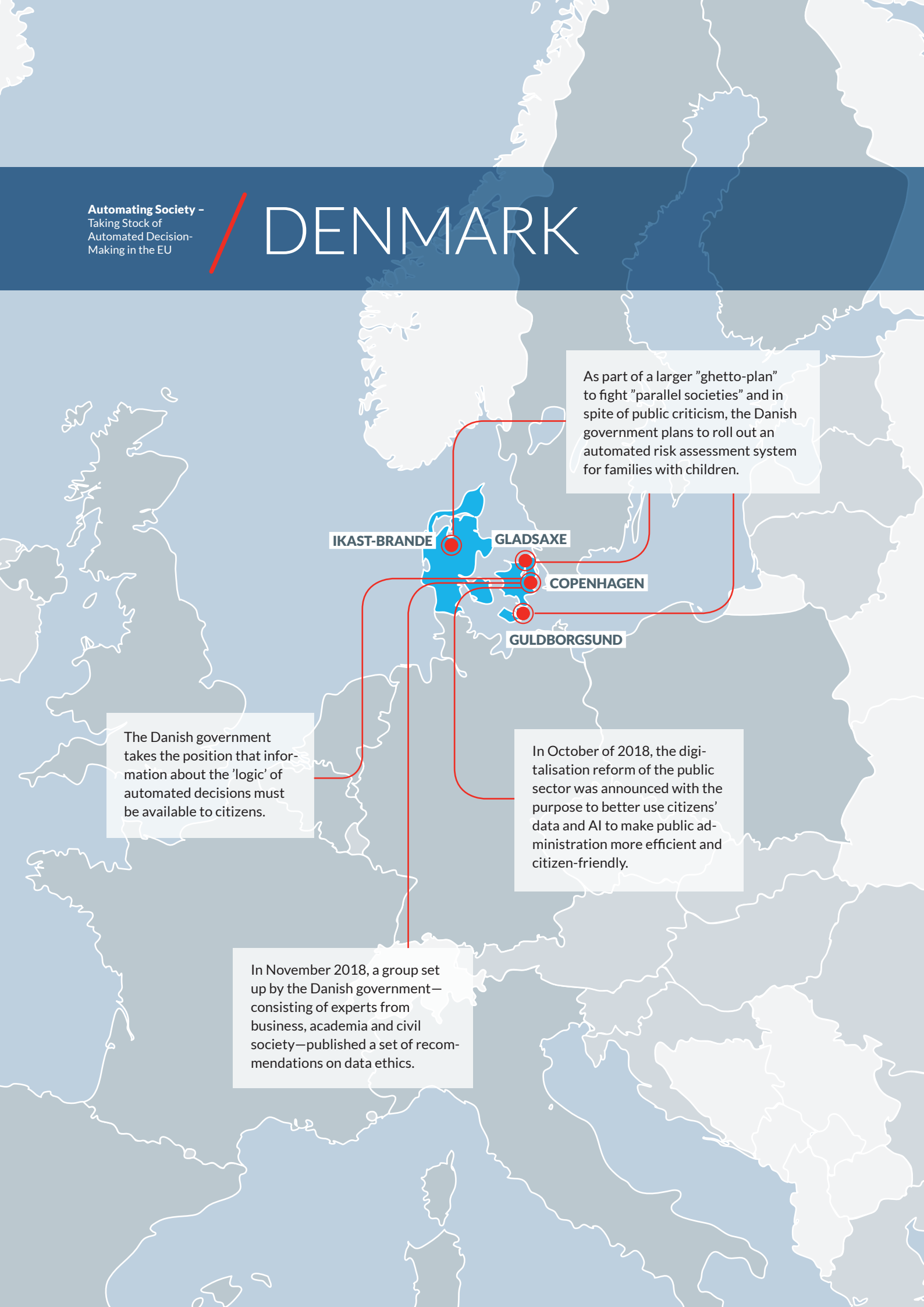
In September 2018, Federal and local police issued a press release to say that they have big plans for predictive policing and already see the possibility that, from the next legislature (after the council elections of October 14, 2018), predictive policing experiments can begin in Antwerp and other local police zones. According to the spokesperson of the Federal Police, they are still working on the tools and building the systems. The data that will be used for the analyses will come from the police databases, for instance the frequency that certain crimes appear in certain areas. In addition, data from external sources will also be important. Predictive policing is mostly seen as a tool to help the police do their work more efficiently. ⧉ **[BE 25]**

Rosamunde van Brakel

is a postdoctoral researcher at the Law, Science, Technology & Society research group at the Vrije Universiteit Brussel. She finalised her PhD Dissertation in 2018 on Taming the future: A rhizomatic analysis of preemptive surveillance of children and its consequences. In addition to this she is executive director of Privacy Salon and managing director of the annual international Computers, Privacy and Data Protection (CPDP) conference.

/ DENMARK

As part of a larger "ghetto-plan" to fight "parallel societies" and in spite of public criticism, the Danish government plans to roll out an automated risk assessment system for families with children.

IKAST-BRANDE

GLADSAXE

COPENHAGEN

GULDBORGSUND

The Danish government takes the position that information about the 'logic' of automated decisions must be available to citizens.

In October of 2018, the digitalisation reform of the public sector was announced with the purpose to better use citizens' data and AI to make public administration more efficient and citizen-friendly.

In November 2018, a group set up by the Danish government—consisting of experts from business, academia and civil society—published a set of recommendations on data ethics.

DENMARK

BY BRIGITTE ALFTER

The Danish government states that it wants to actively further the development of Artificial Intelligence (AI) and related education. The focus is clearly on the potential for economic growth. Activities include support for digital qualifications in general education, funding for research and support for business innovation. A government commission on data ethics has recommended labelling products and services that contain AI-based technology, and it also suggests the creation of a permanent data ethics council. Digital tools, AI, and automated decision-making (ADM) systems are being integrated into public administration processes. Many such ADM systems are not discussed by the wider public and are simply considered to be efficient administration. But some cases have led to widespread political debate, such as a surveillance and early warning system for children in vulnerable circumstances. Specialists in civil society and academia call for transparency, accountability, and adjusted legislation that balances both digital efficiency and civil liberties.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ Digitalisation reform of the public sector

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

On October 23 2018, the digitalisation reform of the public sector was announced. Its purpose was to better use citizens' data, and to use AI to make public administration more efficient and citizen-friendly. ⬈ **[DK 1]** The plan includes a strategy for the public sector, plans for a data ethics council, and the ability for each citizen to see all the data held about him or her, including a log about who accessed this data and when. The plan also points towards providing private companies—such as insurance and banks—access to public data, including citizens' data, and thus support the national strategy for digital growth passed earlier the same year. The plan is supported by an investment fund of 410 million Danish Crowns (€55 million) for 2018-2022. ⬈ **[DK 2]** Denmark is rather unique in that, since the 1960s, data compiled on citizens has been recorded with a unique identifier, a personal number. The government hopes to use this data when developing Artificial Intelligence based on large data sets. ⬈ **[DK 3]** Though the term automatisation is only mentioned in connection to 'routine' tasks, one of the explicit purposes of the plan is to use Artificial Intelligence to better service citizens in areas such as medical prediction systems, or better control of fraud. The plan is set to be implemented over the coming years.

/ National strategy for digital growth

A national strategy for Denmark's digital growth ⬈ **[DK 4]** was published in January 2018 with the overall purpose of stimulating growth. The logic is business driven: that digitalisation leads to improved productivity per worker, that about one third of jobs in Denmark can potentially be automated, and that digital developments create new competition for Danish business. The Danish government and two other political parties have allocated 1 billion Danish Crowns (€134 million) from November 2017 until 2025 for stimulation and development efforts. In addition, certain tax incentives are offered for new initiatives and development in the field. The activities include a Digital Hub ⬈ **[DK 5]** set up in a public-private-partnership by three ministries and three business organisations:

the chamber of commerce, the industry association, and the financial sector. The purpose is to help matchmaking between specialised research, competence and investment, to do further research and to market Denmark as a digitally attractive business environment. Other elements are: a Technology Pact ⬈ **[DK 6]** that allocates resources to digital education in schools (this is the single largest sum budgeted) and includes coding for teachers and children or case-oriented projects in collaboration between schools and local businesses ⬈ **[DK 7]**; a focus on the use of data as a growth driver, including open public data and public-private data sharing; revision of existing regulation to make it easier for business to develop and use new technologies; a catalogue of legislation to be adjusted to the needs of business and consumers and the strengthening of IT security in business. This strategy has to be considered within the context of automated decision-making, even though the strategy does not explicitly mention ADM.

/ Strategy on digital health

In January 2018, Denmark passed a strategy on digital health for 2018-2022 ⬈ **[DK 8].** The plan is to develop a number of digital or digitally supported services, including automatisation, prediction and 'decision support'.

The government's ethics council, Etisk Råd, advocates a balanced approach with a particular focus on privacy rights. ⬈ **[DK 9]** Doctors' organisations are generally in favour of digital developments, though they raise the flag when they see the physician-patient privilege threatened, or when terminology lacks legal clarity. On a more general level, digital developments—including in the health sector and particularly when it comes to (automated) predictions—are addressed by data ethic advocates/consultants, ⬈ **[DK 10]** who, for example, point out the dangers of insecurity related to automated predictions.

/ Government strategy for research and education

The government strategy for research and education from December 2017 ⬈ **[DK 11]** emphasises the use of digital technologies. This covers both the development of entirely new technologies and the application of digital technologies in business and in the public sector. Among other initiatives, the government is also working towards a national centre for digital technologies.

/ Public authorities' digitalisation strategy

In May 2016, public authorities at the national, regional and local level—including administrative bodies such as ministries as well as implementing bodies like public hospitals, schools etc.—decided on a digitalisation strategy 2016-2020. ⬈ **[DK 12]** However, this strategy did not explicitly include automated decision-making.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

A number of fora in Denmark discuss digital growth, opportunities, needs and ethics. Most of them address this on a general level, however some address automated decision-making in particular.

/ Siri-Commission

The Siri-Commission was initiated by a social liberal politician and the union of engineers. Its leading group ☑ [DK 13] consists of high level trade union and business representatives predominantly. Its purpose ☑ [DK 14] is to look into growth and job opportunities connected to Artificial Intelligence and to raise awareness of the effects of such changes on Danish society. In September 2011, the Siri Commission published a report prepared by a data ethics consultancy with a number of recommendations, including for example the requirement that there should always be humans to have the last word, for privacy and data ethics to be built into any design by default, to fight data bias, to use AI in an inclusive way, and to develop standards on how to explain algorithms. ☑ [DK 15]

/ Think tank DataEthics

The think tank DataEthics was founded in 2015 by four women with backgrounds in law, journalism, and business. DataEthics.eu pushes the ethical questions of digital development including of Artificial Intelligence and automated decisions. ☑ [DK 16]

/ Rule-of-law think tank Justitia

The rule-of-law think tank Justitia has a general focus on rule of law questions, but is aware of digital and automated decision-making considerations and contributes to the public debate in the field. ☑ [DK 17]

REGULATORY AND SELF-REGULATORY MEASURES

/ Government commission on data ethics

In November 2018, a group set up by the Danish government—consisting of experts from business, academia and civil society—published a set of recommendations on data ethics. The report by the Data Ethics Commission ☑ [DK 18] agreed on nine recommendations including a permanent and specialised ethical council.

/ Political agreement on digital ready legislation

In January 2018, an agreement on digital ready legislation ☑ [DK 19] ☑ [DK 20] was aproved by all parties and the relevant guidance came into force in July 2018. ☑ [DK 21] This is set to be renegotiated in 2020 and replaces a previous similar agreement and guidance ☑. [DK 22] The seven principles of the agreement concern 1) clear rules, 2) digital communication, 3) automated administration, 4) shared terminology and reuse of data, 5) safe and secure data handling, 6) use of public infrastructure and 7) prevention of abuse and mistakes.

The independent legal think tank Justitia raised concerns about a number of elements, including automated case management by public authorities. In particular, Justitia flagged the lack of rules in which cases would continue to determine when a human is needed to make a decision, and the lack of transparency in automated decision-making which would allow greater scrutiny. Justitia also raised concerns that digitalisation guidance did not sufficiently take into account the legal security and privacy of citizens. ☑ [DK23]

/ Implementation of the GDPR in Denmark

Denmark implemented the GDPR with legislation that came into force on May 23, 2018. ⤴ **[DK 24]** However, a minority group of four centre-left parties in the Danish parliament were critical of the legislation when it was adopted. They said that the public authorities' right to access and to combine the personal data of citizens was too intrusive, and that it meant that there was no obligation to inform citizens. ⤴ **[DK 25]** ⤴ **[DK 26]** The minority group protested against the "far reaching possibilities" of combining data from different authorities, for example "place of living, nationality, missed doctor's appointments, unemployment, mental illness or drug abuse". The minority group also emphasised the need to inform citizens about this compilation of data about individual citizens.

/ Algorithm transparency in automated decisions

Along with the implementation of the GDPR, Denmark takes the position that information about the 'logic' of automated decisions must be available to citizens: "When the data controller responsible solely has to inform about the 'logic' of the automated decisions, a more detailed description of the basis for the process cannot be demanded. The important thing must be that the affected person can understand the considerations underlying the process and how 'the system' reaches the various decisions". ⤴ **[DK 27]** Legal scholar Hanne Marie Motzfeldt fears that automated decision-making software could be used in public administration, while at the same time there are no control mechanisms in place such as technology that can trace patterns applied by the software.[1] In her analysis of the fulfilment of necessary 'system transparency' she refers to existing guidance by the Danish Ombudsman on the obligation of authorities to fulfil all necessary principles for public administration. These include the 'officialprincip'—the implicit duty to investigate the facts of a case before making a decision. ⤴ **[DK 28]** However, various cases lead her to the—explicitly tentative— conclusion, that administrative law fully secures transparency and control mechanisms.[2] Motzfeldt observes a lack of a "lively and qualified debate" about the legal aspects on the balance between technical possibilities and a wish for more efficient administration on the one hand, and the price to be paid by citizens' trust, including demands for full transparency, on the other.[3] Motzfeldt is the leader of the newly established Centre for Law and Digitalisation at the University of Århus. ⤴ **[DK 29]**

ADM IN ACTION

/ ADM in the Danish administration

Automated decisions are applied in Danish administration, often without much ado. For example, student stipends for higher education are decided by combining the student's online application with the information that he or she is accepted to undertake a course of education that qualifies for such a stipend, and the funds are then transferred to the bank account of the student. This process is based upon the law on study stipends ⤴ **[DK 30]** and attracts little attention. But other fields of ADM do attract attention.

---

1   Motzfeldt, H. M. (2018). Retssikkerheden bør følge med den automatiserede forvaltning. I R.F. Jørgensen, & B. K. Olsen (red.), Eksponeret: Grænser for privatliv i en digital tid (S. 227-243). Gad, p. 238.

2   Motzfeldt, H. M., p. 240.

3   Motzfeldt, H. M., p. 242.

/ Banks and insurance

Profiling and automated decisions are present in the banking and insurance sector in Denmark. These activities are regulated via data protection laws and overseen by the Danish data protection authority, Data Tilsynet. In a more general comment on the dilemmas of the information economy, Rikke Frank Jørgensen, a specialised human rights expert, points at information imbalances, as well as the need to address the "fundamental discrepancies between a personalised information economy on one side and a society based upon respect for privacy and data protection".[4]

Credit scoring

Following media coverage in 2005, Data Tilsynet produced a precedent decision for Experian, an international credit scoring company working in Denmark. The decision addressed the parameters for the credit prediction of individuals and companies using scoring systems called 'Consumer Delphi (individuals)' and 'Commercial Delphi (companies)'. According to the company, the parameters used to make the decision included birth date, address and address changes, and open or closed registrations in a debtors' register, including the size of the registered debt. In its decision, Data Tilsynet discussed the parameters included and described which parameters should be included and emphasised. In the case of an individual complaint, the decisive parameters for a given decision must be disclosed. ☒ [DK 31] Over the years, Data Tilsynet has repeatedly addressed company permissions, questions from Parliament, and complaints by consumers concerning their credit scoring practices. ☒ [DK 32]

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

Public discussion about the matter is not widespread, though specialised media focusing on digital and computer developments do pick up on the question. Media outlets have, for example, described how banks are capable of targeting customers based on their consumption patterns ☒ [DK 33], or by using interviews with companies offering profiles and predictions. ☒ [DK 34]

Car insurance

Car insurers offer rebates if drivers install a box ☒ [DK 35] to measure speed, acceleration, deceleration and g-force. The company offers a fixed 25% rebate for installing the box. Another company at some point pondered building a mobile app ☒ [DK 36] that included driving instructions and measurements leading to a quarterly, monthly, or potentially even more frequent adjustments to the car insurance premiums. This app, however, is currently unavailable.

/ Children in vulnerable circumstances – tracing model as part of the 'ghetto plan'

An article in the Politiken national newspaper published at the beginning of 2018 caused a public uproar. Three local authorities had asked for exemption from the usual data protection rules to run an experiment ☒ [DK 37] ☒ [DK 38] to trace children with special needs from a very early stage. The model was named Gladsaxe, after the municipality in the suburbs of Copenhagen. The two other municipalities involved were Guldborgsund and Ikast-Brande, representing a rural, and a mixed rural-industrial community. The purpose was to trace children who were vulnerable due to social circumstances even before they

---

4　　Jørgensen, R. F. (2018). Når informationsøkonomien bliver personlig (When Information Economy gets personal). In R. F. Jørgensen, & B. K. Olsen (red.), Eksponeret: Grænser for privatliv i en digital tid, Gad., p, 86

showed actual symptoms of special needs. Based on previous use of statistics, the authorities decided to combine information about 'risk indicators'.

The model used a points-based system, with parameters such as mental illness (3000 points), unemployment (500 points), missing a doctor's appointment (1000 points) or dentist's appointment (300 points). Divorce was also included in the risk estimation, which was then rolled out to all families with children. ⎘ **[DK 39]** ⎘ **[DK 40]** After the story about this system was published in Politiken—along with the government's apparent plans to roll out the model all over Denmark—the public reacted strongly. The notion of a points-based system reached far and wide. Many refer to it—in jokes and irony—on a colloquial basis, such as "Oh no, I forgot the dentist. As a single parent I'd better watch out now…". In addition, an evaluation scheme of children's well-being and development at kindergarten was unveiled. Individual evaluations were prepared and stored without the knowledge of parents and in breach of existing legislation. ⎘ **[DK 41]** While the latter is data gathering rather than automated flagging—and thus only creates material that can potentially be used for automated risk assessment—the public and political reactions to this scheme were strong, including the reaction from academia. ⎘ **[DK 42]**

In spite of the public criticism ⎘ **[DK 43]**, the Danish government planned to roll out the early tracing model from Gladsaxe to the whole country. This is part of a larger 'ghetto-plan' to fight 'parallel societies'. It is a plan that sets a number of criteria for a neighbourhood to qualify as a 'ghetto' and then introduces a series of special measures, such as higher punishments for crimes, forcing children into public day care at an early age, lifting the protection of tenants in order to privatise public housing, tearing down entire building blocks and—indeed—applying the automated risk assessment system for families with children. ⎘ **[DK 44]** ⎘ **[DK 45]** In September 2018 the minister responsible mentioned a planned legal act [5], but by December 2018 the speaker on legal affairs of the government coalition partner Liberal Alliance said to newspaper Politiken that the proposal had been shelved [6].

Other publicly funded, automated risk assessment experiments in the field of social welfare are under development. For example, a project that measures chronically ill patients' behaviour in order to estimate when or how further efforts are necessary. ⎘ **[DK 46]** ⎘ **[DK 47]** ⎘ **[DK 48]** ⎘ **[DK 49]** Significant government funding for investment in this field is allocated for 2018-2022. ⎘ **[DK 50]**[7] Data ethics consultants urge the general public to be mindful of democratic control, privacy, and ethical questions with such projects. ⎘ **[DK 51]**

/ EFI – the failed tax collection system

EFI (short for one shared collection system, Et Fælles Inddrivelsessystem, 2005 - 2015) was initiated in 2005 to create a digital collection system for taxes at the local as well as at the national level. Before EFI, these taxes were collected separately and de-centrally. The new system had serious technical as well as legal flaws and led to the loss of billions of crowns for the public, due to expired or uncollected claims. It was halted in 2015. ⎘ **[DK 52]** An official investigation found mistakes that could have led to the illegal collection of tax,

---

5   https://www.ft.dk/samling/20171/almdel/sou/spm/558/svar/1507910/1933577.pdf (amended after editorial deadline)

6   https://politiken.dk/indland/art6919255/Regeringen-har-lagt-sin-plan-om-overv%C3%A5gning-af-b%C3%B8rnefamilier-i-skuffen (amended after editorial deadline)

7   410 million Danish Crowns to be invested in the field from 2018 to 2022, press release by Danish government from October 2018 ⎘ **[DK 51]**

wrong registration, or expiration of claims. [DK 53] According to legal expert Hanne Marie Motzfeldt, mistakes in EFI's "data, design, programming and integration in the administrative bodies led to administration in conflict with the law"[8]. One of the problems was lack of insight into the processes: "Precise knowledge about the functioning of data and business processes that were 'cast' into the IT systems were largely placed with the IT provider" and not with the authority itself. Further "data and systems often were so badly documented that [the tax authority] did not have sufficient insights into them". [DK 54]

/ Predictive policing

In the autumn of 2016, public tender documents and FOI requests obtained by journalists showed that the Danish police and the Danish police intelligence service had ordered a digital system from the US company Palantir. [DK 55] Tender documents showed that the system should be able to handle and make searchable very different data sources. These include document and case handling systems, investigation support systems, forensic and mobile forensic systems, as well as different types of acquiring systems such as open source acquisition, and information exchange between external police bodies. [DK 56] [DK 57] In that context, experts voiced criticism that this was a portent to making 'predictive policing' possible. [DK 58] The new digital system for Danish police and Danish police intelligence was adopted as part of anti-terrorism measures. [DK 59]

Two years previously, in 2014, an automatic license plate control system was introduced by Danish police. Using this system, police cars with a camera mounted at the front could automatically screen license plates, check them against several databases, and then indicate on a screen in the police car if there was a match alleging an offence. Human rights specialists have raised questions about the scale of surveillance. [DK 60]

/ Profiling and price adjustment

The Danish Consumers' Council—a prominent, independent consumers' association—explicitly warns the public of price discrimination. "Address, cookies and other personal information can be used to adjust individual prices on goods, so consumers do not pay the same price when shopping. This is unfair and makes the market opaque". [DK 61] While not referring to individual cases, the group provides instructions to the public to avoid profiling via cookies. [DK 62]

/ Public sector data – Planning elderly care and HR document control

The Municipality of Copenhagen—responsible for very different tasks stretching from technical infrastructure and schools to social security and care of the elderly—cooperates with three universities in the capital region to use public sector data and develop automated procedures. [DK 63]

To improve the planning of care for the elderly, the municipality hoped to predict the needs of individuals. Data already logged about assistance, hospitalisation and from semi-structured text by caretakers were aggregated and combined to create an individual history. By analysing three months back in time, it was possible to predict with 80 percent

---

8   Motzfeldt, H. M., p. 231.

precision when significantly more care would be needed, the municipality claimed. The logging and analysis did not change the need, but allowed for more targeted assistance and planning.

Another project concerned the human resources department, where the task was to control whether all relevant information about individual employees was obtained and correctly filed. This includes documents such as contracts, work permits and criminal records. The automated solution included a script to find the documents, place them in a cloud platform, read them with OCR translation, and use a self-made algorithm to find the relevant documents. The automated process was said to be 90 percent accurate due to bad scans of some documents, while manual checks were estimated to be 95 percent accurate. The cost of running the automated scan was estimated at 7,000 Danish Krone (just below €1,000), compared to an estimated manual workload of three months for ten people at 1,000,000 Danish Krone (or €134,000). ⧉ [DK 64]

/ Udbetaling Danmark – automated payments and control of social funds

In Denmark, pensions, child allowances, unemployment support and many other social welfare payments are made by one centralised body called Udbetaling Danmark. This body has far-reaching access to data on individuals from a wide range of sources, which is regulated by the Law on Udbetaling Danmark. ⧉ [DK 65] Data about a citizen from local municipalities, unemployment savings agencies and so forth are used to select a sample of cases for further control. In the first three quarters of 2017, this led to a selection of samples of just above two percent, a quarter of which was taken to further detailed control[9]. An analysis by Birgitte Arent Eiriksson, deputy director of the legal think tank Jusititia, relates the level of respect for privacy to the quality of decisions as estimated by a public control body, and reaches the conclusion, that "efficiency and surveillance" are rated higher than rights and the rule of law. Eiriksson's report asks for a deeper analysis including the origin and treatment of the data[10], or in other words that transparency and proportionality need to be addressed.

/ IBM Watson & breast screening

In 2017, the Capital Region of Denmark entered into an agreement with IBM ⧉ [DK 66] to test at least two AI projects per year using the company's Watson system. Watson—marketed by name and with humanoid terminology such as a "new colleague who does not drink coffee" — was set to be used for routine preventive mammography screenings at two hospitals in the region. One of the arguments for using Watson was that there was a lack of qualified doctors who specialised in radiography. ⧉ [DK 67] The new agreement was made in spite of reports about difficulties during a previous test with another tool, Watson Oncology, which according to media reports recommended life-threatening medication to cancer patients. ⧉ [DK 68] While the evaluation of working with Watson Oncology was positive overall—and future interest was indicated by IBM, the health services and universities—the difficulties with the project were described as needing "further development and adaptation before the technology can be implemented for clinical use. For example, doctors and

---

9   Eiriksson, B. A. (2018). Social digital kontrol er på kant med borgernes ret til privatliv. I R. F. Jørgensen, & B. K. Olsen (red.), Eksponeret: Grænser for privatliv i en digital tid (S. 227-243). Gad., p. 34.

10   Eiriksson, B. A., p. 38 ff.

Watson only agreed on 27% of treatment suggestions. This was likely due, among other factors, to the fact that the system used had been trained in the US following American guidelines and practices". ☑ [DK 69] In connection with these agreements, academics at the IT University of Copenhagen warned against being "deceived" by the new technology and called for better information of the public and of decision makers. ☑ [DK 70]

On a more general level, and endorsing the new technologies to further good health[11], professor of health and law, Mette Hartlev of Copenhagen University suggests that fundamentally new legislation is needed in the field of health and data to counteract discrimination, inequality, breaches of privacy, data security and so forth.

---

11   Hartlev, M. (2018). Sundhedsdata sætter patienters privatliv under pres. I R. F. Jørgensen, & B. K. Olsen (red.), Eksponeret: Grænser for privatliv i en digital tid (S. 227-243). Gad.

Brigitte Alfter

is an award-winning Danish-German journalist specialising in European affairs. After a career in local and national Danish media, including a position as Brussels correspondent, she turned to cross-border collaborative journalism with ad hoc teams as well as with existing networks such as the ICIJ. She is the co-founder of several cross-border journalism projects and structures such as the Farmsubsidy.org project, the Scoop-project, Journalismfund.eu, the European Investigative Journalism Conference & Dataharvest, the Investigate Europe team and the European Journalism ARENA. Brigitte has published a book about cross-border collaborative journalism in Danish and German. An English version will be published in 2019.

/ FINLAND
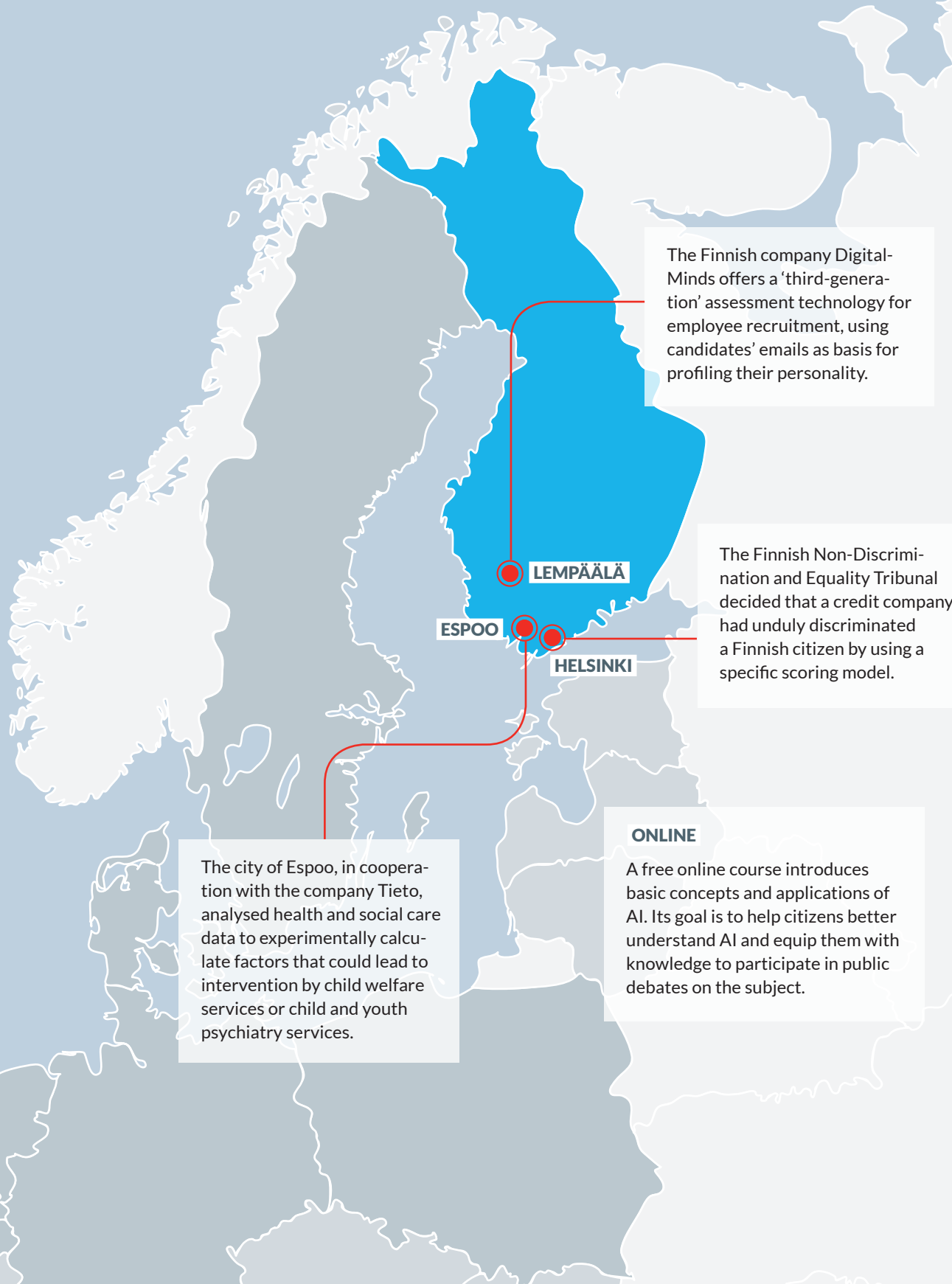
The Finnish company Digital-Minds offers a 'third-generation' assessment technology for employee recruitment, using candidates' emails as basis for profiling their personality.

LEMPÄÄLÄ

ESPOO

HELSINKI

The Finnish Non-Discrimination and Equality Tribunal decided that a credit company had unduly discriminated a Finnish citizen by using a specific scoring model.

The city of Espoo, in cooperation with the company Tieto, analysed health and social care data to experimentally calculate factors that could lead to intervention by child welfare services or child and youth psychiatry services.

ONLINE

A free online course introduces basic concepts and applications of AI. Its goal is to help citizens better understand AI and equip them with knowledge to participate in public debates on the subject.

FINLAND

BY MINNA RUCKENSTEIN AND JULIA VELKOVA

At the level of Finnish central government, automation is predominantly discussed in relation to Artificial Intelligence. The government's goal is to pool together societal-wide resources to foster developments around AI and automated decision-making. The governmental AI Programme consists of concrete initiatives to boost economic growth and to revitalise companies and the public sector. An ethical information policy—commissioned to address questions of data ownership and the effects of automation on Finnish society—will be finalised soon. Civil society actors, in turn, are concerned with ethical uses of data that underpin automated decision-making. A key actor in the process is an international non-governmental organization, MyData Global, which grew out of a Finnish data activism initiative. Other issues on the agenda are the discriminatory nature of credit scoring and clarifications of mistakes made by automated processes related to tax assessments. Such clarifications are important in the light of ongoing automation projects in the public sector. Social insurance institutions, among others, are struggling to resolve the challenges of ensuring compatibility with existing laws, and with the difficulties in justifying the logic of automated decision-making processes. The start-up sector for ADM solutions is growing rapidly and involves work on a variety of projects including prospective employee personality assessments, health diagnosis, and content moderation of online discussions.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ The Artificial Intelligence Programme

The Artificial Intelligence Programme ⬀ **[FI 1]**, commissioned by the Minister of Economic Affairs, was launched in May 2017 and the final AI strategy is due in April 2019. To implement the work of the initiative, a steering group was established, chaired by Pekka Ala-Pietilä, CEO and co-founder of Blyk, and former president of Nokia Corporation and Nokia Mobile Phones. The group's first report, published in October 2017, introduced AI as an opportunity for Finland to be among the winners of the economic transformation that is believed will revolutionise industries and organizations from transport and healthcare to energy and higher education.

The AI Programme describes the application of AI as a societal-wide pressure for rapid transformation that offers opportunities for economic growth and renewal for companies and the public sector, if the opportunities are dealt with in a systematic manner. One of the stated aims of the programme is the formation of "a broad-based consensus" of the possibilities of AI to foster "a good artificial intelligence society". A broad consensus on the matter is seen as essential, because of the limited resources of a small nation. Therefore, the contributions to the AI field must be implemented efficiently with a consistent emphasis on economic impact.

While specifics of automated decision-making are not mentioned in the AI programme, the simultaneous improvement of service quality and the level of personalization, alongside expected efficiency gains, point to expectations of considerable automation in the provision of public services. The AI programme foresees efficiency increases in service provision as

the main benefit for the public sector, in particular when it comes to the provision of health-care services. With the help of AI, services provided by the public administration become free of the confines of time and location. Digital services can utilise appropriate information at the right time, and hence proactive interventions can be made to enhance citizen wellbeing. In short, AI is expected to help the public sector to predict service needs, and respond in a timely manner to each citizen's needs and personal circumstances.

Key points of Finnish AI strategy

The proposed actions for the Finnish AI strategy emphasise the competitiveness of companies through the use of AI ⧉ **[FI 2]**. This goal is supported by various proposals, for instance, by ensuring that Finnish data resources are accumulated and enriched systematically, their technical and semantic interoperability is ensured and that datasets are utilised in all sectors of society. The adoption of AI is simplified with "platform economy trials" that bring together companies and research facilities for the piloting of AI solutions. Such piloting is supported by an independent facilitator—the CSC IT Center for Science, a non-profit organization owned by the state (70%) and higher education institutions (30%)—which offers computational and data storage resources.

In terms of public administration, the AI strategy promotes public-private partnerships and the renewal of public services. In addition, new cooperation models are being established to boost digital service development. The goal is to build public services that are the best in the world, always available and in any language needed. AI applications are developed to better anticipate and predict service needs of the future. The strategy promises that time consuming queues and telephone appointments will be eliminated by the use of personalised services and digital assistants. Work towards that aim, however, has only just begun.

/ Elements of Artificial Intelligence

The strategy work emphasises that Finns must have access to AI literacy—a guaranteed basic understanding of AI principles. In order to support this goal, an English-language online course called 'Elements of Artificial Intelligence' ⧉ **[FI 3]** was developed by the Department of Computer Science at the University of Helsinki in partnership with the technology company Reaktor to form part of the Finnish AI Programme. The two parties involved produced the course work pro bono. The course introduces basic concepts and applications of AI and machine learning with the aim of increasing public understanding of AI and better equipping people to participate in public debates on the subject. The societal implications of AI, such as algorithmic bias and de-anonymisation, are introduced to underscore the need for policies and regulations to guarantee that society can adapt well to the changes the ever-widening use of AI brings. The course is free for anyone to attend, and close to hundred thousand participants have already signed up.

/ FCAI – Finnish Center for Artificial Intelligence

The Finnish AI strategy emphasises the necessity for a Center of Excellence for AI and applied basic research that will attract top-level international expertise. The recently established Finnish Center for Artificial Intelligence (FCAI) ⧉ **[FI 4]** aims to respond to such a need by becoming a nationwide cross-disciplinary competence centre for AI, initiated by Aalto University, University of Helsinki, and VTT Technical Research Centre of Finland. The mission is to create "Real AI for Real People in the Real World", a type of AI which operates in collaboration with humans in various everyday domains.

As part of the Center, the FCAI Society ☑ [FI 5], a group which consists of experts from philosophy, ethics, sociology, legal studies, psychology and art, will explore the impact AI has in all aspects of our lives. Both FCAI Society and FCAI researchers are committed to engaging in public dialogue when considering the wider implications of AI research.

/ Request of clarification about using automated tax assessments

In September 2018, the Deputy Parliamentary Ombudsman, Maija Sakslin, requested clarification from the Tax Administration about using automated tax assessments. ☑ [FI 6] In 2017, the Deputy Ombudsman had made two complaints regarding mistakes made by automated processes. In her position paper, the ombudsman writes that in two of her complaint settlements, the Tax Administration reported that taxation procedures of self-assessed taxes are mostly automated. Although the mistakes made by automated processes had been fixed, the ombudsman is concerned about how taxpayers' legal protection, good administration and accountability of public servants are secured in automated taxation decisions. ☑ [FI 7]

The ombudsman states that it is problematic that the requests for information letters and taxation decision letters sent by the automated taxation systems do not include any specific contact information, only general service numbers of the Tax Administration. Moreover, the automated taxation process produces no justification concerning the decisions it makes. If there are problems with the decision, the issue is handled at a call centre by a public servant who has not taken part in the decision-making and has no detailed information about the decision. Based on the two complaints, the ombudsman stated that, in terms of automated taxation, taxpayers' legal rights to receive an accurate service and justification of the taxation decisions are currently unclear. The Deputy-Ombudsman has requested the Ministry of Finance to obtain the reports needed from the Tax Administration and the Ministry to produce their own report regarding the legality of automated taxation. The report was due in November 2018.

/ Report on ethical information policy in an age of Artificial Intelligence

In March 2018, the Ministry of Finance set up a group to prepare a report on ethical questions concerning information policy and AI. A draft of the report was made publicly available and open to comment until the end of October and is currently being finalised. The report discusses policies regarding individuals' rights to their own data (MyData), openness of AI solutions, and the prevention of adverse effects stemming from automated decision-making on support systems and society. The policy needs are described on a very general level, and any possible regulatory or legislative measures resulting from the report will take place under the next government. ☑ [FI 8]

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ MyData Global

In October 2018, an international association for the ethical use of personal data, MyData Global ☑ [FI 9], was founded to consolidate and promote the MyData initiative that started

in Finland some years ago. The association's purpose is to empower individuals by improving their right to self-determination regarding their personal data. The human-centric paradigm is aimed at a fair, sustainable, and prosperous digital society, where the sharing of personal data is based on trust as well as a balanced and fair relationship between individuals and organisations. The founding of this association formalises a network of experts and stakeholders, working on issues around personal data uses, and they have been gathering in Helsinki annually since August 2016.

As machine learning and AI-based systems rely on personal data generated by and about individuals, the ethics of these technologies have been at the forefront in the meetings of the MyData community. Much like personal data itself, the potential of these technologies for individual as well as collective and societal good is recognised as being enormous. At the same time, considerations of privacy and the rights of individuals and collectives to know what data about them is being used, and for what purposes, as well as to opt out selectively or wholesale, must be taken seriously, argue the proponents of the MyData model. [FI 10] At the core of MyData is an attempt to bring experts together to find arenas of appropriate intervention in terms of building an ethically more robust society.

REGULATORY AND SELF-REGULATORY MEASURES

/ Ethical Challenge for Enterprises

In terms of the ethics, Finland's AI Programme challenges enterprises to share their understanding and use of ethical principles with the aim of making Finland a model country for the ethical uses of AI. Questions that this challenge [FI 11] is addressing include the following: Is the data used to train AI biased or discriminating? Who is responsible for the decisions made by AI? This work started recently—the kick-off event for companies was held in October 2018. The aim is to promote self-regulation in companies and formulate principles that define how AI could be used in fair, more transparent and trust-building ways. Leading Finnish enterprises that were the first to join the ethics challenge include the K Group, OP Group and Stora Enso and currently more than fifty companies are participating in the challenge. [FI 12]

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/automating-society*

OVERSIGHT MECHANISMS

/ Non-Discrimination and Equality Tribunal

In April 2018, the National Non-Discrimination and Equality Tribunal prohibited a financial company, specialising in credits, the use of certain statistical methods in credit scoring decisions. The Non-Discrimination Ombudsman had requested the tribunal to investigate whether the credit institution company Svea Ekonomi AB was guilty of discrimination in a case that occurred in July 2015. The case considered whether the company did not grant a loan to a person in connection to the purchase of building materials online. [FI 13]

Having received a rejection to the loan application, the credit applicant, a Finnish-speaking man in his thirties from a sparsely populated rural area of Finland, asked the company to justify the negative decision. The company first responded by saying that their decision required no justification, and then that the decision had been based on a credit rating made by credit scoring service using statistical methods. Such services do not take the creditwor-

thiness of individual credit applicants into account and therefore, the assessments made may significantly differ from the profile of the individual credit applicant. This, the credit company agreed, may seem unfair to a credit applicant.

The credit applicant petitioned the Non-Discrimination Ombudsman who then investigated the case for over a year. The credit rejection in question was based on data obtained from the internal records of Svea Ekonomi, the credit company, information from the credit data file, and the score from the scoring system from an external service provider. Since the applicant had no prior payment deficits in the internal records of the credit company, nor in the credit data file, the scoring system gave him a score based on factors such as his place of residence, gender, age and mother tongue. The company did not investigate the applicant's income or financial situation, and neither was this information required on the credit application. As men have more payment failures than women, men are awarded fewer points in the scoring system than women and similarly, those with Finnish as their first language receive fewer points than Swedish-speaking Finns. Had the applicant been a woman, or Swedish-speaking, he would have met the company criteria for the loan.

After failing to reconcile the case, the Non-Discrimination Ombudsman brought the case to a tribunal. The ombudsman decided that the credit company was guilty of discrimination based on the Non-Discrimination Act. The applicant's age, male gender, Finnish as the mother tongue and the place of residence in a rural area were all factors that contributed to a case of multiple discriminations, resulting in a decision not to grant a loan. Discrimination based on such factors is prohibited in section 8 of the Non-Discrimination Act and in the section 8 of the Act on Equality between Women and Men. The tribunal noted that it was remarkable that the applicant would have been granted the loan if he were a woman or spoke Swedish as his mother tongue.

The National Non-Discrimination and Equality Tribunal prohibited Svea Ekonomi from continuing their discriminatory practices and imposed a conditional fine of €100,000 to enforce the prohibitive decision.

ADM IN ACTION

/ Benefit processes at the Social Insurance Institution of Finland (Kela)

Kela is responsible for settling benefits under national social security programmes. Around forty core benefits are handled by Kela, including health insurance, state pensions, unemployment benefits, disability benefits, child benefits and childcare allowances, student financial aid, housing allowances, and basic social assistance. While benefits are an important income source for underprivileged groups in particular, most Finns regardless of income level or social status receive benefits from Kela during their lifetime. Kela hands out benefits of approximately €15.5 billion annually. [FI 14]

While more traditional automated information systems have been used for decision automation in Kela for decades, AI, machine learning and software robotics are seen as an integral part of their future ICT systems. [FI 15] Ongoing and potential AI developments include chatbots [FI 16] [FI 17] for customer service, automated benefit processing, detection (or prevention) of fraud or misunderstanding, and customer data analytics.

In the case of benefit processing in general, and automated processes in particular, it is required that procedures must conform to existing legislation. From an automation perspective, the relevant phases of the benefit process include submitting the application, prerequisite checks, and the making of the decision. In the application phase, the system can provide information and advice on benefits. However, benefit legislation can allow for various combinations of benefits to be applied in a given situation and because benefits are interlinked, different combinations may produce different total benefits for the citizen. Therefore, a parameter for automation is how the system should prioritise benefits. Deciding which is the 'best' combination of benefits means deciding what outcome should be prioritised—the client's total received benefits, or some other goal (political, economic or otherwise).

The automation of prerequisite checks entails, first, checking whether the information provided is sufficient for decision-making, valid, and trustworthy; and second, whether other benefits affect, or are affected by the applied benefit. Once these checks have been completed, the decision can be made. In most cases decisions are based on a regulated set of rules, and machine learning can be taught using validated data Kela already has from previous decisions.

A problem that Kela faces as a public organization is how to communicate the results and the reasoning behind the decision-making process to citizens. In the case of automated decision-making, decisions are essentially probabilistic in nature, and models are based on features of similar historical cases. How can decision trees be translated into text that is understandable to the customer? And how is the probabilistic accuracy of automated decision-making communicated? In addition to accuracy, one relevant parameter for the customer is response time— this can be measured in weeks when humans process applications, but can be practically instantaneous when an automated system is used.

Legislative processes that produce regulation that decisions are based on complicate the automation of the benefit process. Kela's decisions are based on more than 200 pieces of separate regulation, owned by 6 ministries, and spanning 30 years. Separate laws may be semantically incompatible which complicates their translation into code: for example, different regulations contain more than 20 interpretations of 'income'. In addition, benefit laws change, and automated models need to behave in a new way after the new law comes into force. When new benefits are created, no decision data is available, which means machine learning needs to be taught using, for example, proxy data created specifically for this purpose. This points towards new employee roles that automated decision-making creates in an organization like Kela: in addition to data scientists who supervise and train the systems and analyse data, data validators evaluate decisions made by the systems and refine models based on evaluation, and data creators produce new data to teach the models. ⧉ **[FI 18]**

/ Child welfare and psychiatry services – the use of predictive AI analytics

In an experiment undertaken by the City of Espoo, in cooperation with software and service company Tieto, AI was used to analyse anonymised health care and social care data of Espoo's population and client data of early childhood education. The goal of the experiment, carried out in 2017 and 2018, was to screen service paths from the data by grouping together risk factors that could lead to the need for child welfare services or child and youth psychiatry services. ⧉ **[FI 19]** Tieto and Espoo processed data from the years 2002 to 2016,

consisting of 520,000 people who had used the services during that time, and 37 million client contacts. The experiment produced preliminary results about factors that could lead to the need for child welfare services or child and youth psychiatry services. For example, the AI system found approximately 280 factors that could anticipate the need for child welfare services. ⤤ **[FI 20]**

The experiment was the first of its kind in Finland. Before it, data from public services had never been combined and analysed so extensively in Finland using AI technologies. By combining data from several sources, it was possible to review public service customer paths by families, as the city's data systems are usually individual-oriented. The City of Espoo is planning to continue experimenting with AI. The next step is to consider how to utilise AI to allocate services in a preventive manner and to identify relevant partners to cooperate with towards that aim. The technical possibilities to use an AI-based system to alert health and social care personnel to clients' cumulative risk factors have also been tested. Yet, Espoo states that ethical guidance regarding the use of an AI system like this is needed. For example, issues of privacy and automated decision-making should be carefully considered before introducing alert systems for child welfare services. Among other things, Espoo is now discussing ethical questions related to the use of such alert systems with expert organisations such as The Finnish Center for Artificial Intelligence (FCAI). As a partner in the experiment, Tieto has gained enough knowledge to develop their own commercial AI platform for its customers in both the public and private sector which raises further ethical questions.

/ DigitalMinds – Assessing workers' personality based on automated analysis of digital footprints

The Finnish start-up company DigitalMinds ⤤ **[FI 21]** is building a 'third-generation' assessment technology for employee recruitment. Key clients (currently between 10 and 20 Finnish companies) are large corporations and private companies with high volumes of job applicants. Personality assessment technologies have been used since the 1940s in job recruitment. At first, these came in the form of paper personality tests that were filled in by prospective job candidates to assess their personality traits. Since the 1990s, such tests have been done in online environments.

With their new service, DigitalMinds aims to eliminate the human participation in the process, in order to make the personality assessment process 'faster' and 'more reliable', according to the company. Since 2017 it has used public interfaces of social media (Twitter and Facebook) and email (Gmail and Microsoft Office 365) to analyse the entire corpus of an individuals' online presence. This results in a personality assessment that a prospective employer can use to assess a prospective employee. Measures that are tracked include how active individuals are online and how they react to posts/emails. Such techniques are sometimes complemented with automated video analysis to analyse personality in verbal communication (see, for example, the HireVue software ⤤ **[FI 22]**). The results produced are similar to the ones made by traditional assessment providers, i.e. whether a person is an introvert/extrovert, attentive to details etc.

The companies which use this technology do not store any data, and they are required by Finnish law to ask the prospective job candidates for informed consent to gain access to their social media profiles and email accounts in order to perform the personality test. The candidates use a Google or Microsoft login to access the DigitalMinds platform and then they input their credentials themselves, avoiding direct disclosure of these credentials to the company. According to DigitalMinds, there have been no objections to such analysis so

far from prospective candidates. A crucial ethical issue that must be considered, however, is around the actual degree of choice that a candidate has to decline access to her personal and former/current work email accounts, as well as to social media profiles in online accounts—if a candidate is in a vulnerable position and needs a job, they might be reluctant to decline such an assessment of data that may be very personal, or include professional/trade secrets within emails.

In addition, the varied amount of data that is available online about individuals makes comparison between candidates difficult. According to the company, the trustworthiness of the data is not a big issue, if there is a sufficient corpus available online. The company also waives responsibility by suggesting that they do not make actual decisions, but that they automate the process of assessment based on which decisions are made. An important issue that this case raises is the degree to which individuals' online digital/information in social media and emails should be considered trustworthy. It potentially harms disadvantaged groups who may have reasons to have concealed or fake online personalities.

/ Digital symptom checkers

A strategic project of the current governmental programme for "self-care and digital value services", ODA, has developed the OmaOlo ("MyFeel" in English) service which includes digital symptom checkers. ☑ **[FI 23]** Such checkers are now available for lower back pain, respiratory infection symptoms and urinary tract infections and there will be many more in the future. ☑ **[FI 24]** The symptom checkers are based on The Finnish Medical Society Duodecim's medical database and evidence-based clinical practices' Current Care Guidelines.

In practice, the symptom checker asks simple yes-or-no-questions about patients' symptoms and then it offers advice about whether the condition needs medical attention or if self-care is sufficient. The checkers also include questions about acute symptoms that need immediate care. If the patient answers "yes" to any acute symptoms, the system recommends that they contact emergency services as soon as possible. The OmaOlo service has a disclaimer emphasizing that the aim of the symptom checker is not to produce a diagnosis but to make an evaluation of care needs. This means that it is possible that the assessments made by the symptom checkers are not always accurate, or that they are interpreted as more precise than they actually are. The questionnaires and the algorithms used to analyse them are based on research, when such research is available, or on collective medical care experience.

The symptom checkers have been tested in several areas of Finland. In testing, OmaOlo tried to collect as much feedback of the services as possible for further development and validation. Preliminary observations suggest that compared to the assessments made by health care personnel, more patients than before are referred to medical care by symptom checkers. The checkers, and other OmaOlo services, will be introduced more generally to the public in 2019. A new state-owned company, SoteDigi, aims to produce and develop digital health and social care services and will be in charge of the development and the distribution of the OmaOlo services.

/ Utopia Analytics – Automated Content Moderation on Social Media and e-commerce websites

Utopia Analytics ☑ **[FI 25]** is a Finnish startup that specialises in automating content moderation on social media, public discussion forums, and e-commerce websites. The company

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

brands itself as working with the aim to "bring democracy back into social media" by keeping "your audiences talking, your online communities safe and your brand authentic". The service that they offer is a self-learning AI (neural network) which analyses the corpus of data on a given website, and then starts moderating content. At first, moderation is done alongside human moderators who test and 'train' the decisions of the system. After some time, human moderators take the role of supervisors of the algorithm, and only control its decisions in controversial or less straightforward cases.

The service has been used on the Swiss peer-to-peer sales service site tutti.ch where it moderates sales advertisements for inappropriate or inaccurate content in several languages, as well as on the largest Finnish online forum, Suomi24. Regarding moderation on Suomi24, the company reports a quantitative increase in the number of moderated posts—from 8,000 to 28,000—and a related increase in advertisements on the website. It is not clear what kind of content is being moderated in each case and how. It seems that such decisions are contextual and can be set up within a framework of each platform that implements the algorithm.

Minna Ruckenstein

Minna Ruckenstein works as an associate professor at the Consumer Society Research Centre and the Helsinki Center for Digital Humanities, University of Helsinki. Ruckenstein has studied human-technology involvements from various perspectives, exploring how people use direct-to-consumer genetic testing, engage with self-tracking technologies and understand algorithmic systems as part of daily lives. The disciplinary underpinnings of her work range from anthropology, science and technology studies to consumer economics. She has published widely in top-quality international journals, including Big Data & Society and Information, Communication and Society. Prior to her academic work, Ruckenstein worked as a journalist and an independent consultant, and the professional experience has shaped the way she works, in a participatory mode, in interdisciplinary groups and with stakeholders involved. Most recent collaborative projects explore social imaginaries of data activism.

Julia Velkova

Julia Velkova is a digital media scholar and post-doctoral researcher at the Consumer Society Research Centre at the University of Helsinki. Her research lies at the crossing between infrastructure studies, science and technology studies and cultural studies of new and digital media. She currently works on a project which explores the waste economies behind the production of 'the cloud' with focus on the residual heat, temporalities and spaces that are created in the process of data centres being established in the Nordic countries. Other themes that she currently works with include algorithmic and metric cultures. She is also the Vice-Chair of the section "Media Industries and Cultural Production" of the European Communication and Research Education Association (ECREA). Her work has been published in journals such as New Media & Society, Big Data & Society, and International Journal of Cultural Studies, among others.

FRANCE

Automated processing of traffic offences became a massive revenue stream—and the government agency responsible, ANTAI, ignores legal requirements to disclose their algorithms.

Obliged by law since 2017, the French administration has to release algorithms upon request and prohibits solely automated systems that are not fully explainable to the subject of the decision.

RENNES

PARIS

BORDEAUX

MARSEILLE

Journalists at Next INpact use freedom of information requests to test transparency laws and opened up some algorithms used by the state.

The think tank FING (New Generation Internet Foundation) advocates for a measure of "mediation": Users have to be offered support, whether automated or human, whenever needed, when confronted with automated decision-making.

FRANCE

BY NICOLAS KAYSER-BRIL

French lawmakers' first confrontation with ADM came in 1978, when they introduced a law to ban it unconditionally. The new law granted citizens the right to access and modify their personal data and required corporations and the government to request pre-authorisation before creating databases of people. At the same time, it prohibited credit scores—something which continues to this day.

Over time, it appears that the police felt only loosely bound by this law. According to a 2011 parliamentary review of police databases, over a quarter of them had not been legally authorised. [FR 1] They have since been legalised [1].

Recent legal changes have followed the same pattern. Following a law change in 2016, it became mandatory for all branches of government to make their algorithms transparent. However, journalists who reviewed three ministries discovered that none of them had complied with the regulation.

During the state of emergency, under President Hollande (2015-2017), laws were passed to allow the police to undertake automated mass-surveillance of internet activity. In addition, the Ministry of the Interior is currently looking at ways to link all CCTV footage to biometric files and automate the detection of 'unusual' behaviour. [FR 3] However, several watchdogs—some financed by the state—lobby for more transparency relating to how ADM is used.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ AI strategy of the government

The French government's current AI strategy follows the 'AI for Humanity' [FR 4] report, written by MP and mathematician Cédric Villani in March 2018, and adopted by President Macron the following month. [FR 5] However, the president chose not to take up most of the report's recommendations, including a proposal to double the salaries for young academics working in AI.

Instead, Macron chose to adhere to the following guidelines: To focus the strategy on health, transport, security and the environment, and to double the number of students in AI (although no target date was set for this), and to act at the European level. In addition, he wants to simplify regulations involving AI experiments and create a €10 billion innovation fund—part of which would pay for work in AI.

---

1   A 2018 parliamentary report showed that the chaotic situation prior to 2018 has been fixed. However, at least 11 databases containing personal information—mostly compiled by the intelligence agencies—are not controlled by any organisation. It is unclear as to whether these databases are already being used to detect devious behaviour. However, the military plans to go down this route. See [FR 2]

In March 2017, President Hollande's administration published its own AI strategy, but it was coordinated by a different branch of the government to the one Macron used. ⧉ **[FR 6]** This multiplicity of AI reports will not help the French administration to know exactly what the current official strategy is.

/ National Digital Council

In 2011, President Sarkozy's administration created a National Digital Council (*Conseil National du Numérique* or CNNum) which continues to this day. Council members are selected by the government and only have consultative powers. Historically, the CNNum has included high-profile entrepreneurs and think tank employees whose collective Twitter clout gives them an over-sized influence.

By its nature, the CNNum is resolutely pro-business, but it has also clashed with the government on issues concerning civil liberties. In particular, it opposed ADM with regard to state surveillance of citizens (see the Regulatory Measures, Mass Surveillance section for further details). In that case, it argued that predictive algorithms would reinforce social biases. ⧉ **[FR 7]**

The CNNum did not write its own AI report, but it did contribute in large part to the government's first AI strategy. This strategy highlighted the need for well-balanced human-computer interactions, and co-operation between all stakeholders before the deployment of ADM in companies.

After a clash with the government in December 2017—when the Minister for Digital Affairs refused to let one expert join the college—the CNNum resigned en masse. ⧉ **[FR 8]** As a result, the government nominated a new set of experts who were more aligned with its opinions. Some commentators think that this episode might reduce the influence and relevance of the consultative body in the future.

/ Etalab – Entrepreneurs in the public interest

Etalab, the open data outlet of the French government, runs a programme that lets highly educated, tech-savvy youths work as "entrepreneurs in the public interest". This means that they are embedded in the administration for a ten-month period to experiment with new ways of doing things and some of them have worked on ADM. ⧉ **[FR 9]** However, the impact of this programme is hard to assess. For instance, one entrepreneur—who was embedded in the fraud-tracking service of the Ministry of Finance—said that his algorithm would not replace the legacy system. At this stage, it is unclear whether these public-sector entrepreneur posts will amount to anything more than glorified internships.

POLITICAL DEBATE ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ AlgoGlitch

In late 2017, the CNNum (see above) tasked the Medialab at Science-Po, a research centre closely associated with digital activism, to study how algorithmic glitches 'trouble' the general public. AlgoGlitch ⧉ **[FR 10]**, which folded in August 2018, mostly ran workshops on the topic and its results remain purely qualitative.

/ Algo Transparency

Algo Transparency is an informal team of technologists who monitor recommended videos on YouTube—not just French content—and publish the results on the AlgoTransparency website. [FR 11] Funding and network support come, in part, from Data For Good [FR 12], which is a French incubator of civic-tech projects financed partly by private sector foundations.

/ FING

The New Generation Internet Foundation (Fondation Internet Nouvelle Génération, or FING) [FR 13] is a think tank, founded in 2000, which brings together large and medium-sized companies as well as public bodies. It organises conferences and runs the Internet Actu news website, which translates English language articles on ADM and publishes them in French.

On the issue of ADM (which it refers to as "systems"), FING advocates strongly for a measure of "mediation" which it defines as the ability for users to receive support, whether automated or human, whenever needed.

/ La Data en Clair

La Data en Clair (Data in the Clear) [FR 14] started in June 2018 and is an online magazine focused on the ethical aspects of Artificial Intelligence. It published several articles in June, suspended operations for a while, and resumed work in November 2018. [FR 15]

/ La Quadrature du Net

On the NGO side, La Quadrature du Net [FR 16] is dedicated to online freedom, but it only follows ADM issues at the French or European level when they encroach on individual freedoms such as smart cities, online censorship or predictive policing. However, this site does not consider ADM to be a key area of interest and it has never led any campaigns (e.g. petitions, public awareness campaigns) on the issue. La Quadrature du Net tends to focus on other topics, although it has fought on issues related to ADM indirectly. For example, it started a legal battle to oppose the creation of a biometric database of all French nationals—something which is a prerequisite for large-scale face recognition—but La Quadrature du Net ended up losing that battle. [FR 17]

/ NEXT INpact

The online news outlet, NEXT INpact [FR 18] is one of the very few French newsrooms to consistently follow issues related to ADM on the national and governmental level. Their journalists regularly use freedom of information requests to test transparency laws and they were responsible for the publication of some algorithms used by the state. However, they do not cover the private sector with the same intensity.

REGULATORY AND SELF-REGULATORY MEASURES

/ Algorithm transparency

Article L311-3-1 of the legal code of relations between the administration and the public [FR 19] states that any decision made using an algorithm must mention the fact that an

algorithm was used. In addition, the rules defining the algorithm and what it does must be released upon request.

In June 2018, the French Supreme Court for administrative matters (Conseil d'Etat) stated that a decision based solely on an algorithmic system could only be legal if the algorithm and its inner workings could be explained entirely to the person affected by the decision. If this is not possible (because of national security concerns, for instance), then algorithmic decision-making cannot be used. ⤢ **[FR 20]**

This piece of legislation was passed in 2016 and became law on September 1, 2017. A year later, journalists tried to find cases where this law had been applied, but they could not find any. ⤢ **[FR 21]** The law authorising the Parcoursup system (see the ADM in Action section for details) contains a clause that freed it from complying with this regulation. ⤢ **[FR 22]**

/ Autonomous driving

A law currently being debated in parliament aims to clarify responsibilities around autonomous driving. ⤢ **[FR 23]** The stated goal is to allow tests of fully autonomous vehicles. In May 2018, the government announced that the necessary legislation will not be passed until 2022.

When signed into law, it will allow autonomous vehicles up to SAE (Society of Automotive Engineers) level 4 on all roads (high automation— i.e. where the system is in complete control of the vehicle and a human presence is not required). However, its application will be limited to specific conditions. ⤢ **[FR 24]**

Up until now, car manufacturers have used ad-hoc authorisation to test their vehicles. ⤢ **[FR 25]**

/ Computers and Freedom Law from 1978

The Computers and Freedom Law followed an outcry after a government agency planned to connect several databases and use social security numbers as unique identifiers. The result of the argument was the creation of the legal concept of "personal data"—and its protection—in France. It has been modified substantially over the years to keep up with technological changes, most notably in 2004 and in 2017 (in accordance with the GDPR).

Strict limitations on ADM

Article 2 of the Computer and Freedom Law ⤢ **[FR 26]** states that no judicial decision "regarding human behaviour" can be made on the basis of an ADM that "gave a definition of the profile or personality" of the individual. The same applies to administrative and corporate decisions (this disposition now constitutes article 10 ⤢ **[FR 27]** of the law, in a slightly watered-down version that contains several exceptions.).

Pre-authorisation

The law states, in articles 15 and 16 ⤢ **[FR 28]**, that any algorithmic decision-making must be submitted to the CNIL (see under Oversight Mechanisms below). This same law allows, in chapter IV, anyone to access or request modification of his or her personal data stored by any administration or company (this now constitutes article 22 ff. ⤢ **[FR 29]** However, mandatory declaration, which had replaced pre-authorisation a few months after the law was passed, was dropped in 2017).

⤢ **LINKS:** *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

/ Mass surveillance

A 2015 law relating to domestic intelligence allows the police and intelligence agencies to request that internet service providers deploy algorithmic systems to detect terrorist threats using "connection data" only (Article L851-3 of the Code for domestic security [FR 30]).

These deployments are regulated by a consultative body known as the National Commission for the Control of Intelligence Gathering Techniques (*Commission Nationale de Contrôle des Techniques de Renseignement*).

The Ministry of the Interior is required to produce a parliamentary technical report on the use of this measure before end of June 2020. Based on this, lawmakers will decide whether to renew this legal clause. [FR 31]

OVERSIGHT MECHANISMS

/ CNIL – France's data protection authority

The 2016 Digital Republic Law states that the National Commission on Computers and Freedom (CNIL), a nominally independent institution financed by the Prime Minister's Office, must concentrate its energies on the fast-changing digital environment.

There followed a series of forty-five seminars, discussions and debates which resulted in an 80-pages report [FR 32] published in late 2017. The report attempted to explain the issues under scrutiny, especially regarding ADM. Among the recommendations, which included better ethics education, the report also proposed the creation of a national algorithm audit platform. However, this particular recommendation has yet to be discussed in parliament.[2]

ADM IN ACTION

/ Automated processing of traffic offences

Automated processing of traffic offences started in 2003 with the deployment of about 50 radars across France. These now number approximately 4,500 and record over 17 million offences per year—and the trend is increasing.[3] The gross income from these radars is probably slightly over €1 billion per year, making it one of the most visible ADM processes French citizens are likely to encounter.

In 2011, the Ministry of the Interior created a new government agency to manage radars and traffic offences—the National Agency for the Automated Processing of Offences (ANTAI). Somewhat revealingly, ANTAI does not mention anywhere in its annual report how it complies with article L311-3-1 of the legal code covering relations between the

---

2   According to a search of nosdeputes.fr and nossenateurs.fr—two websites that allow users to search transcribed parliamentary discussions.

3   Observatoire national interministériel de la sécurité routière, "La sécurité routière en France : Bilan de l'accidentalité de l'année 2017" [FR 33], 2018, p. 10 8, and "ANTAI, Rapport d'Activité 2017" [FR 34], July 2018.

civil service and the public. This relates back to the judgement, mentioned earlier, which requires transparency around the use of algorithms (see Algorithm transparency in the Regulatory Measures section for details). ANTAI did not answer a request for comment on this point. ☑ **[FR 34]**

/ Bob-Emploi – Matching unemployed people with jobs

In 2016, a 22-year-old caused a media sensation by claiming that he could use algorithms to match jobseekers with job vacancies and thereby reduce unemployment by 10%. As a result of his claim, he received €2.5m from public and private donors ☑ **[FR 35]** to finance a team of nine people. They also gained access to anonymised job seeker data and support from the national employment agency, a rare privilege. ☑ **[FR 36]**

However, the initiative, called Bob-Emploi (Bob-job), failed to reduce unemployment (as of June 2018, only 140,000 users had created an account). ☑ **[FR 37]**

/ 'Black boxes' and mass surveillance

In 2017, some two years after the law that authorised 'black boxes' was passed ☑ **[FR 38]**, the first one was deployed. This consisted an algorithm—placed at the level of internet service providers—and used by the intelligence services to analyse internet traffic in order to detect terrorist threats (see legal dispositions above).

In its annual report, the oversight body for domestic surveillance (which only has consultative powers) stated that it was asked for an opinion on this system, but it provided no information regarding the scope or goals of the 'black box'. ☑ **[FR 39]**

/ Credit scoring

The sale of an individual's credit score is forbidden in France.[4] Credit scoring can only be done within a bank and no investigation can be carried out to find out how the scores are calculated.[5]

CNIL prevents the full automation of credit scoring. One ruling demanded that companies must perform manual checks before someone's name is added to a list of people not paying their dues. ☑ **[FR 40]** A file of people who have defaulted on their credit is maintained by the national bank.[6]

In 2012, the government pushed for the creation of a database containing all loans between €200 and €75,000, together with the names of the debtors. Ostensibly, this was to allow lenders to check whether a debtor was already heavily indebted. However, the law was nullified on privacy grounds by the supreme court in 2013.[7]

---

4   Lazarus, Jeanne. "Prévoir la défaillance de crédit: l'ambition du scoring." Raisons politiques 4 (2012): 103-118. On page 107, the author mentions that the CNIL prohibits the sale of individual credit scores.

5   We were unable to find any journalistic or parliamentary work on the topic. There are, however, companies such as Synapse who sell such services (installing a credit scoring mechanism, as opposed to selling scores).

6   The database is called "Fichier des incidents de remboursement des crédits aux particuliers"

7   AFP/CBanque Loi Hamon: "le Conseil constitutionnel rejette la création d'un fichier des crédits à la consommation", March 13, 2014 ☑ **[FR 41]**. The Supreme Court's ruling was worded in such a way that made experts assume that a file containing the credit histories of non-defaulting clients would never be possible under current legal conditions. See Cazenave, Frédéric, "Surendettement: le fichier positif définitivement enterré", Le Monde, June 25, 2015. ☑ **[FR 42]**

/ Dossier Médical – Personalised health files

Health is one of the four priority areas of the government's AI strategy (see AI strategy of the government in the POLITICAL DEBATES ON AUTOMATION section for details).

In a 2018 report ⧉ [FR 43], Inserm, a government outlet specialising in medical research, stated that the health sector did not use ADM as a whole. However, it added that a few test projects are underway locally or will start in 2019—one of these relates to decision support for the treatment of breast cancer while the other is concerned with complex ultrasound diagnosis.

This lack of activity highlights the failure of a project started in 2004 aimed at digitizing and centralising all personal health information in a "personalised health file" (DMP for Dossier Médical Personnel). The DMP project was supposed to allow ADM to operate in the health sector.[8] The DMP was criticised by the French court of auditors in 2012 as ill-conceived and very expensive. ⧉ [FR 45] As a result, it was rebranded as the "shared health file" (Dossier Médical Partagé). However, it is barely used (just 600,000 DMPs were created in 2017 ⧉ [FR 46]).

Before this, in 2008, pharmacists started building their own central file of individual drug purchases and some 35 million people are now registered on the service. ⧉ [FR 47] The data is used by pharmacists to check possible nefarious drug interactions. Laboratories can also access the database to perform drug recalls. However, the data cannot be sold to third parties.

/ Parcoursup – Selection of university students

In 2018, the French government pushed forward a reform whereby universities had the right to turn down applications from prospective students (previously, anyone with a high-school diploma had the right to enrol). The reform was enacted with the launch of an online tool called Parcoursup, which matched the wishes of students with available offerings.

On top of the drawbacks that plague many industrial-scale projects (numerous experts advised that it would be better to spread this reform over a longer period), Parcoursup was criticised for the opacity of its decision-making process.

When drafting the law that created Parcoursup, the government made sure to protect it from the legal obligation to tell users that decisions were made algorithmically and to make the algorithm transparent. Instead, the government published the source code of the matching section of the platform but remained silent on the sorting part.

It is believed that the sorting part uses personal data to help universities select the students who best fit certain courses of study. ⧉ [FR 48]

Photo: Marion Kotlarski

Nicolas Kayser-Bril

Nicolas Kayser-Bril, 32, is a Berlin-based, French-German data-driven journalist. He created the data journalism team at OWNI, a French news startup, in 2010, then co-founded and led the data journalism agency Journalism++ from 2011 to 2017. He now splits his time between teaching programming at HMKW (Berlin), helping fellow journalists in their data-driven investigations, consulting for various organizations and writing history books.

8   The rationale of the personalised health file was made clear in a parliamentary report ⧉ [FR 44] on the topic in January 2008, among others p. 28.

/ GERMANY

The Alliance of Consumer Protection Agencies published a set of requirements that automated decision-making processes need to fulfil in order to be in line with consumer protection.

The city of Mannheim launched an "intelligent video surveillance" project based on motion pattern recognition.

BERLIN

The German AI strategy states that government wants to find ways to "provide effective protection against bias, discrimination, manipulation or other abusive uses, in particular when using algorithm-based prediction and decision systems."

MANNHEIM

ULM

Several German Data Protection Agencies published a joint position on transparency of algorithms, demanding that very sensitive systems should require authorisation by a public agency.

GERMANY

BY KRISTINA PENNER AND LOUISA WELL

In Germany, a number of commissions, expert groups, platforms and organisations were set up by government and parliament to assess the consequence of a wider application of Artificial Intelligence. While it is to be welcomed that different stakeholders engage in the debate, this profusion introduces the risk that expert groups end up working in parallel on overlapping issues. In addition, some of these groups have very short timelines to produce outcomes on a wide range of questions. Also, some of these outcomes are supposed to feed into government policy, but the process to accomplish this is at times unclear, especially because there is no specific framework for cooperation, neither to develop joint strategies nor to prepare implementations.

Other stakeholder groups are actively working in the field of automated decision-making (ADM). Business associations, academic societies, and a number of civil society organisations have already conducted research and published analyses, discussion and position papers on ADM. There seems to be a consensus that the developments offer a lot of opportunity to benefit individuals and society, but that there is a lot of work needed to keep the risks in check.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ German AI Strategy

The national AI strategy [DE 1], which was developed under the joint leadership of the Federal Ministries of Education and Research, Economics and Energy, and Labour and Social Affairs, was presented at the "Digital Summit" in early December 2018 in Nuremberg.

As part of the federal government's implementation strategy for digitisation ("Hightech-Strategie 2025"), the AI strategy is to be further developed and adapted at the beginning of 2020 "according to the state of discussion and requirements".

The foreword summarises the strategy's goals as follows:

> "The strategy is based on the claim to embed a technology as profound as Artificial Intelligence, which may also be used in sensitive areas of life, ethically, legally, culturally and institutionally in such a way that fundamental social values and individual rights are preserved and the technology serves society and mankind."

At the heart of the document is the plan to make "AI made in Germany" a "globally recognised seal of quality". The strategy focuses on three goals:

- To make Germany and Europe a leading location for the development and application of AI technologies and thus contribute to securing Germany's future competitiveness
- To ensure the responsible development and use of AI for the common good
- To embed AI ethically, legally, culturally and institutionally into society, based on a broad dialogue within society and an active role of politics in shaping the issue

In order to develop Artificial Intelligence in a way that is responsible and focused on the public good, the strategy announces a range of measures. The federal government

- plans to create an observatory and vows to campaign for similar observatories on the EU and international level
- wants to organise a European and transatlantic dialogue focusing on a human-centric development of AI in the workplace
- vows to safeguard labour rights when AI is deployed and used in the workplace
- announces it will actively shape the ethical, legislative, cultural and institutional embedding of AI, facilitating a broad dialogue within society
- wants to develop guidelines for the design and use of AI systems in compliance with data protection law in round-table consultations with data protection supervisory authorities and business associations
- says it will promote the development of innovative applications that support autonomy, social and cultural participation, and the protection of citizens' privacy
- announces the creation of a funding scheme to educate citizens and to support the public-minded design of technology
- says it will develop the platform "Lernende Systeme" into an AI platform to organise a dialogue between politics, academia, business and civil society
- pledges to find a European approach to data-based business models that reflect the common values, and their economic and social structure
- aims to reflect whether the regulatory framework needs to be further developed to secure a high level of legal certainty, and asserts it will promote and require the respect of ethical and legal principles throughout the process of AI development and application

The government "sees itself under the obligation" to support the development of „AI made in Germany" as "a technology for the good and benefit of state and society". By promoting diverse AI applications with the aim to achieve "tangible social progress in the interest of the citizens" it aims to focus on the benefits for humans and the environment and to continue an intensive exchange with all stakeholders. It also points to AI as "a collective national task", which requires collective action and cross-sectoral measures. To achieve these goals, the strategy also pleads for stronger cooperation within Europe, especially when it comes to harmonised and ethically high standards for the use of AI technologies.

Implications of particular relevance to the participatory and fairness aspects of automated decision-making—which also have strong effects on the regulatory and policy framework for the application of ADM—can be found in the following paragraphs:

Section 3.1, focusing on *Strengthening Research in Germany and Europe* and section 3.8, *Making data available and facilitating (re-)use* introduce a "systematic approach to technology" that promotes research on methods for "monitoring and traceability of algorithmic prediction and decision-making systems", as well as research on pseudonymisation and anonymisation methods, and on the compilation of synthetic training data. It states that decisions "must be comprehensible so that AI systems can be accepted as 'trusted AI' and meet legal requirements."

In the section about the *labour market* (3.5), questions about the increasing deployment of AI—and ADM—in the processes of hiring, transfers and dismissals, as well as the monitoring of workers' performance or behaviour, are addressed. The authors emphasise that workers' councils theoretically already have the means, within the framework of their labour representation rights, to influence the application of AI in these fields. They recognise,

however, that people need to be empowered to exercise these rights, especially in the light of knowledge asymmetries. Here, the government asserts that it will examine whether an independent employee data protection law can create more legal certainty for the introduction of respective applications in the workplace.

With regard to the objective to *Use AI for public tasks and adapt expertise of the administration* (3.7), the government is hesitant to take a clear stand on the use of AI-based systems. It describes AI as "an an instrument to contribute information to decision-making that cannot be obtained in an adequate timeframe without AI", but clarifies that "police, intelligence and military assessments and decisions based on them will continue to be in the hands of the employees of the authorities." The strategy mentions that "other areas of application are predictive policing (preventive averting of danger), the protection of children and adolescents against sexualised violence on the Internet, and the fight against and prosecution of the dissemination of abuse representations or social media forensics for the creation of personal profiles", but the strategy does not indicate if and to what extent these could be introduced.

Because AI applications will "increasingly contribute to decision-making in everyday life or control it in the background", the government plans to *Review the Regulatory Framework* (3.9) for gaps in relation to algorithmic and AI-based decision-making systems, services and products. If necessary, it vows to adapt regulation to make systems accountable for possible inadmissible discrimination and disadvantage. The strategy describes how "transparency, traceability and verifiability of AI systems" can be established in order to "provide effective protection against bias, discrimination, manipulation or other abusive uses, in particular when using algorithm-based prediction and decision systems."

The government proposes "to examine the establishment or expansion of government agencies and private auditing institutions to monitor algorithmic decisions with the aim of preventing abusive use and discrimination and averting negative social consequences". This is to include the development of auditing standards and impact assessment standards. "This control system should be able to request full disclosure of all elements of the AI/algorithmic decision-making process without the need to disclose trade secrets."

The strategy further develops ideas to review existing regulation and standards and enhance *"standardisation"*, especially in regard to classifications and terminologies of AI and ethical standards in line with the "ethics by design" approach (3.10).

/ Digital Cabinet and Digital Council

The newly created Cabinet Committee on Digitisation discusses questions of AI and examines answers and solutions feeding into the (implementation of the) national strategy. ↗ [DE 2]

The Digital Council, located in the Chancellery, acts as a consultative body, intended to facilitate a close exchange between politics and national and international experts. Appointed by the government, its members come from a variety of sectors: scientists and practitioners, founders of start-ups and established entrepreneurs. Civil society is not represented. The council members are tasked with finding ways to implement the projects to be defined in the AI strategy. ↗ [DE 3]

/ Platform for Artificial Intelligence

The *Plattform Lernende Systeme* (Platform for Artificial Intelligence) was launched by the Federal Ministry of Education and Research in 2017 at the suggestion of acatech (German Academy for Science and Engineering), and will end by 2022. The platform aims to design self-learning systems for the benefit of society. ⧉ **[DE 4]** In different working groups, experts from science, industry, politics and civil society analyse the opportunities and challenges of AI, and develop scenarios and roadmaps to inform decision-makers. ⧉ **[DE 5]** The working group on "IT Security, Privacy, Legal and Ethical Framework" focuses on how AI can be implemented in a way that respects human autonomy, fundamental rights, equal opportunities, and does not discriminate. In general, the group ascertains "how society can actively participate during the implementation of self-learning, and therefore ADM systems". ⧉ **[DE 6]** In the context of the German AI strategy, the platform presented its new interactive information website on actors, strategies and technologies related to AI at the national, EU and international level. ⧉ **[DE 7]**

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ Enquete Commission on Artificial Intelligence

In June 2018, the German Bundestag set up the "Study Commission Artificial Intelligence – Social Responsibility and Economic, Social and Ecological Potential". ⧉ **[DE 8]** ⧉ **[DE 9]** This so-called Enquete Commission is comprised of 19 MPs from all parliamentary parties and 19 external experts appointed by parliamentary groups. ⧉ **[DE 10]** The mandate of the commission is to examine the opportunities as well as the challenges of AI and its effects on individuals, society, economy, labour and the public sector. The commission's task is to develop recommendations for policy makers based on its findings in order to "use the opportunities offered by AI and to minimise the risks." ⧉ **[DE 8]** The sectors of public administration, mobility, health, care, autonomy, ageing, education, defence, environment, climate and consumer protection will be at the centre of its inquiries. Furthermore, the commission will analyse the impact of AI on democratic processes, gender equality, privacy protection, economy, and labour rights. One of the crucial issues of the commission's work will be the assessment of regulation for scenarios in which automated decisions about individuals and society are prepared, recommended or made. The final report is due in 2020. ⧉ **[DE 11]**

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ AlgorithmWatch

The NGO AlgorithmWatch ⧉ **[DE 12]** argues that automated decision-making systems are never neutral but shaped by economic incentives, values of the developers, legal frameworks, political debates and power structures beyond the mere software developing process. They campaign for more intelligibility of ADM systems and more democratic control. The organisation's mission is to analyse and watch ADM processes, explain how they work to a broad audience and to engage with decision-makers to develop strategies for the beneficial implementation of ADM. At the moment, AlgorithmWatch is looking at a German credit scoring algorithm, the automation of human resources management, and is mapping the growth of automation across a variety of sectors.

/ Alliance of Consumer Protection Agencies

Bundesverband der Verbraucherzentralen (alliance of consumer protection agencies ⧉ **[DE 13]**) is funded in large part by government grants. It advocates for consumer protection in a wide range of fields, including e-commerce, social networks and networked gadgets. In a detailed position paper published in late 2017 ⧉ **[DE 14]**, the alliance drafted a set of requirements that it believes automated decision-making processes need to fulfil in order to be in line with consumer protection laws. These requirements include comprehensibility of ADM processes, redress mechanisms for affected individuals, labelling of processes that involve ADM, adaption of the liability framework, and intensified research into such systems.

/ Bertelsmann Stiftung

In its project Ethics of Algorithms ⧉ **[DE 15]**, a team at Bertelsmann Stiftung analyses the influence of algorithmic decision-making on society. The project's stated goal is to shape algorithmic systems in order to increase social inclusion (Teilhabe) by developing solutions that align technology and society.

/ Bitkom

Bitkom (Association for IT, Telecommunications and New Media ⧉ **[DE 16]**) represents digital companies in Germany. It publishes guidelines and position papers on technological and legislative developments. In a publication they produced on AI, they examine in great detail the economic significance, social challenges and human responsibility in relation to AI and automated decision-making. ADM's ethical, legal and regulatory implications are analysed from the point of view of corporate and social responsibility. ⧉ **[DE 17]** In a separate paper focused on the "responsible use of AI and ADM", the authors recommend that companies develop internal and external guidelines, provide transparency as to where and how ADM is used, conduct a cost-benefit calculation with regard to users, guarantee accuracy of data and document provenance, address and reduce machine bias, and design high-impact ADM systems in a way that a human retains the final decision. ⧉ **[DE 18]**

/ Chaos Computer Club

The Chaos Computer Club (CCC ⧉ **[DE 19]**) is an association of hackers who advocate for more transparency in government, freedom of information, and the human right to communication. The annual conference, Chaos Communication Congress, is one of the most popular events of its kind with around 12,000 participants. In recent years, the conference featured various presentations and discussions ⧉ **[DE 20]** about how automated decision-making processes change society.

/ German Informatics Society

With almost 20,000 private members and about 250 corporate members, Gesellschaft für Informatik (German Informatics Society ⧉ **[DE 21]**) is the largest association for computer science professionals in the German-speaking world. In a recent study, the authors analysed what they call "algorithmic decision-making" with a focus on scoring, ranging from credit scoring to social credit systems. ⧉ **[DE 22]** They recommend intensifying interdisciplinary research into ADM, incorporating ADM in curricula of computer science and law, fostering data literacy in economics and social sciences, developing procedures for the testing and auditing of ADM systems, standardised interfaces, transparency require-

ments, and creating a public agency to oversee ADM. In a response to the government's AI strategy ⬚ **[DE 23]**, the society stated that a European contribution to the development of AI has to be the development of explainable algorithmic decision-making and explainable AI that prevents discrimination.

/ Initiative D21

Initiative D21 ⬚ **[DE 24]** is a network of representatives from industry, politics, science and civil society. Its working group on ethics published discussion papers on ADM in medicine, public administration, elderly care and other sectors. ⬚ **[DE 25]**

/ Netzpolitik.org

The platform Netzpolitik.org ⬚ **[DE 26]** states that its mission is to defend digital rights. Its team of authors cover the debates about how political regulation changes the Internet and how, in turn, the Internet changes politics, public discourse and society. Articles that deal with ADM focus mostly on the issue from a data protection and human rights perspective.

/ Robotics & AI Law Society

As a relatively new actor in the field of AI in Germany, Robotics & AI Law Society (RAILS) states that it wants to address challenges of "(AI) and (intelligent) robotics" and "actively shape the discussion about the current and future national and international legal framework for AI and robotics by identifying the need for regulatory action and developing concrete recommendations." ⬚ **[DE 27]** The organisation also states that its "aim is to ensure that intelligent systems are designed in a responsible way, providing a legal framework that facilitates technical developments, avoids discrimination, ensures equal treatment and transparency, protects fundamental democratic principles and ensures that all parties involved are adequately participating in the economic results of the digitalization." ⬚ **[DE 28]**

/ Stiftung Neue Verantwortung

The think tank Stiftung Neue Verantwortung (Foundation New Responsibility) ⬚ **[DE 29]** develops positions on how German politics can shape technological change in society, the economy, and the public sector. It works on ideas of how algorithms can be used to foster the common good, explore the potential of ADM for fairer decision-making processes, and attempt to create a development process for ADM that respects the common good in its design.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

REGULATORY AND SELF-REGULATORY MEASURES

/ Data Ethics Commission

The Data Ethics Commission was set up by the government to "develop ethical standards and guidelines for the protection of individuals, the preservation of social cohesion and the safeguarding and promotion of prosperity in the information age." ⬚ **[DE 30]** ⬚ **[DE 31]** The commission is asked to propose a "development framework for data policy, the use of algorithms, artificial intelligence and digital innovations" for the government and parliament. The committee consists of 16 experts from politics, academia and industry and is led by the Federal Ministry for Justice and Consumer Protection and the Federal Ministry of the Interior, Building and Community.

One of three key areas addressed in the questions posed to the commission relates to algorithmic decision-making. It emphasises that "people are being evaluated by technical processes in more and more areas of life". ⬚ **[DE 32]** The Commission looks at how algorithms, AI and digital innovations affect the life of citizens, the economy or society as a whole. It also keeps an eye on what ethical and legal questions arise, and it gives advice on which technologies should be introduced in the future and how they should be used. In the first paper of recommendations sent to the government, the commission argued that ethics does not "mean primarily the definition of limits; on the contrary, when ethical considerations are addressed from the start of the development process, they can make a powerful contribution to design, supporting advisable and desirable applications." ⬚ **[DE 33]**

/ Data Protection Agencies' joint position on transparency of algorithms

In a joint position paper ⬚ **[DE 34]**, the data protection agencies of the federal government and eight German federal states stated that greater transparency in the implementation of algorithms in the administration was indispensable for the protection of fundamental rights. The agencies demanded that if automated systems are used in the public sector, it is crucial that processes are intelligible, and can be audited and controlled. In addition, public administration officials have to be able to provide an explanation of the logic of the systems used and the consequences of their use. Self-learning systems must also be accompanied by technical tools to analyse and explain their methods. An audit trail should be created, and the software code should be made available to the administration and, if possible, to the public. According to the position paper, there need to be mechanisms for citizens to demand redress or reversal of decisions, and the processes must not be discriminating. In cases where there is a high risk for citizens, there needs to be a risk assessment done before deployment. Very sensitive systems should require authorisation by a public agency that has yet to be created.

/ Draft-law on automated driving

In 2017, the German government proposed the adoption of a law on automated driving. ⬚ **[DE 35]** The draft law determines that the driver will hold the final responsibility over the car, even if the system was driving automatically. Hence, drivers must always be able to interrupt the automated driving and assume control themselves. The level of automation can be differentiated into partly automated, highly automated, fully automated and autonomous driving. Autonomous driving is not covered by the draft law.

/ Ethics Commission on Automated and Networked Driving

The Federal Ministry of Transport and Digital Infrastructure convened this commission in 2016 to work on ethical questions that arise with the introduction of self-driving cars. ⬚ **[DE 36]** Under the premise that the automation of traffic promises to reduce accidents and increase road safety, the commission discussed an array of questions and in 2017, it published 20 rather abstract principles on automated and networked driving. ⬚ **[DE 37]** These principles include the suggestion that automated systems can only be approved if, on balance, they promise to lead to a reduction of harm, and that a public agency for the assessment of accidents involving automated vehicles should be installed to gather evidence in order to guide future development.

/ German Ethics Council

In order to inform the public and encourage discussion in society, the German Ethics Council ⎆ **[DE 38]** organises public events and provides information about its internal consultations. The Council addresses questions of ethics, society, science, medicine and law, and the consequences for individuals and society related to research and development, in particular in the field of life sciences and their application to humanity. In the past, the German Ethics Council has discussed ethical issues related to the health sector. Those discussions focussed on developments in synthetic biology in 2009, particularly on technological innovations and the implications in the field of genetic diagnosis and its use in medical practice[1] in 2012. ⎆ **[DE 39]** ⎆ **[DE 40]**

In 2017, the Ethics Council published a report on the consequences of automatically collecting, analysing and then inferring insights from large data sets in the health sector. ⎆ **[DE 41]** It concluded that these processes lead to stratification – the algorithmic grouping of individuals based on certain characteristics. This can be useful for diagnostics and therapy, but the council cautions against erroneous attribution of data. "Big Data" in the context of health promises to be especially fruitful in biomedical research, healthcare provision, for insurers and employers, and in commercial applications. However, the council warns against the exploitation of individuals' data, discrimination, the linking of health data to other data by commercial actors, an "excessive regime of self-control aided by such services and devices [that] can contribute to an exaggerated drive for optimisation detrimental to personal health", and privacy breaches.

In 2017, the Council's annual conference focused on "autonomous systems" in general, and discussed "how intelligent machines are changing us". ⎆ **[DE 42]**

/ German Institute for Standardisation

The German Institute for Standardisation (Deutsches Institut für Normung e.V. (DIN) ⎆ **[DE 43]**) works on standardisation for e-mobility, energy transition, Industry 4.0, secure online identification, smart cities, and smart textiles. One working group, established at the beginning of 2018 as part of DIN's Standards Committee on Information Technology and Applications (NIA), focuses on ethical and societal issues concerning AI. ⎆ **[DE 44]** Their first step was to draft a working paper on the terminology of AI itself, a topic considered to be of particular importance. The development process of new norms is closed to the public and further details are not yet available.

ADM IN ACTION

/ Automated management of energy infrastructure

The transition to renewable energies, technological innovation, and the more flexible trade of electrical energy across Europe continue to change and challenge Germany's energy infrastructure. Smart grids and a growing number of private households producing their own electricity (so-called prosumers) only add to the challenge. ADM in the area of distribution grids ensures a more resilient supply and enables a smoother feed-in and control of

---

1   High-throughput procedures of molecular diagnostics (rapid, automated procedures to analyse a large number of samples).

decentralised and fluctuating electricity on different levels. On the transmission grid level, automation has already been in use for some time. It was established to support the complex task of distributing, monitoring and controlling electricity flows to ensure system and frequency stability. Today, the new generation of Supervisory Control and Data Acquisition (SCADA) systems, are an essential part of this infrastructure. They use machine learning and predictive analysis for real time monitoring and controlling in order to avoid frequency fluctuations or a power supply failure. [DE 45] At the same time, government experts point out that energy supply, as a critical infrastructure, is under special threat from cyber attacks. [DE 46] Advocacy groups also warn against extensive surveillance capabilities that come with the implementation of smart meters in private homes. [DE 47]

/ Credit scoring

In Germany, like in many other countries, private credit scoring companies rate citizens' creditworthiness. A low score means banks will not approve a loan or may reject a credit card application, and Internet service providers and phone companies may deny service. This means credit scoring has enormous influence on people's ability to meaningfully participate in everyday life. In Germany, the leading company for scoring individuals, SCHUFA Holding AG, has a market share of up to 90 per cent (depending on the sector), so it wields immense power. The civil society organisations AlgorithmWatch and Open Knowledge Foundation Germany initiated a project called OpenSCHUFA. [DE 48] They asked individuals to donate their SCHUFA scores to the project, so that data journalists and researchers could systematically scrutinise the process. The results of the analysis were published in late November 2018. [DE 49] Data journalists from Spiegel Online and Bayerischer Rundfunk public broadcasting station identified various anomalies in the data. For instance, it was striking that a number of people were rated rather negatively even though SCHUFA had no negative information on them, e.g. on debt defaults. There were also noticeable differences between alternate versions of the SCHUFA scoring system. The credit report agency offers to its clients (such as local banks or telecommunication companies) a score that is specifically tailored to their business segment. In one of the available cases the scores differ by up to 10 per cent between version 2 and 3, depending on the version of the scoring system the client uses as a reference. This is an indicator that the SCHUFA system itself deems its version 2 to be somewhat lacking. However, version 2 is apparently still used by SCHUFA clients.

/ "Intelligent video surveillance system" in the city of Mannheim

The city of Mannheim in Baden-Wuerttemberg launched an "intelligent video surveillance" project, developed in cooperation with the Fraunhofer Institute for Optronics, Systems Engineering and Image Evaluation. The technology is not based on face recognition, but on "automatic image processing". [DE 50]

Installed sequentially, by 2020 around "76 cameras will be used to monitor people in central squares and streets in the city centre and scan their behaviour for certain patterns" [DE 51] that "indicate criminal offences such as hitting, running, kicking, falling, recognised by appropriate algorithms and immediately reported to the police". In this way, the "police can intervene quickly and purposefully and save resources in the future", says the city's website. Critics warn that the application of such behavioural scanners, here in the form of video surveillance with motion pattern recognition, "exerts a strong conformity pressure and at the same time generates many false alarms", as "it is also not transparent to which 'unnatural movements' the algorithms are trained to react. Thus, lawful behaviour,

such as prolonged stays at one place, could be included in the algorithms as suspicious facts." ⤴ **[DE 51]**

The stated goal of the municipal government is to make "the fight against street crime more efficient and provide more security to people in the city ". ⤴ **[DE 50]** The system is one component of a "comprehensive security concept of the City of Mannheim", which also includes a regular "urban security audit", "Security Round Tables", mobile security guards at certain exposed locations, and the promotion of crime prevention by the organisation 'Security in Mannheim'.

Kristina Penner

is the Executive Advisor at AlgorithmWatch. Her research interests include ADM in social welfare systems, social scoring and the societal impacts of ADM as well as the development of participatory and empowering concepts. Her analysis of the EU border management system builds on her previous experience in research and counselling on the implementation of European and German asylum law. Further experience includes projects on the use of media in civil society and conflict sensitive journalism, as well as stakeholder involvement in peace processes in the Philippines. She holds a master's degree in International Studies / Peace and Conflict Research from Goethe University in Frankfurt.

Louisa Well

is a Master's student at the University of Edinburgh, studying Science and Technology in Society. She wrote her BA thesis on Internet censorship in Turkey, and graduated in Political Science and English Literature from the University of Heidelberg. Freedom online, Internet governance, surveillance and algorithmic accountability are her favourite topics of research. In 2018, she did an internship at AlgorithmWatch.

The Università Cattolica del Sacro Cuore in Milan designed the "Abbiamo i numeri giusti" ("We have the right numbers") system to help health institutions pick the most efficient and effective treatments for patients automatically, while at the same time optimising public spending.

The Nexa Research Center for Internet and Society looks at unexpected and unintended consequences "arising from the increasing complexity and autonomy of some algorithms" and focuses on the "responsibilities of designers and data scientists".

The "eSecurity" project is based on the idea that "in any urban environment, crime and deviance concentrate in some areas (streets, squares, etc.), and that past victimization predicts future victimization". It is supposed to provide complex automated assistance to law enforcement agencies.

TRENTO

MILAN

BOLOGNA

TURIN

ROME

"RiskER" is used to automatically predict the risk of hospitalization. It is used in 25 "Case della Salute" (health centres) in Emilia-Romagna.

In its whitepaper on AI, the government-appointed Digital Transformation Team demands that "predictions of impact and measurement of the social and economic effects of AI systems" are provided in order to "enhance the positive effects and reduce risks".

ITALY

BY FABIO CHIUSI

In recent years, discussions in Italy about automation have increased both at the institutional level and by the public at large. Academic institutions, mainstream media and civil society have started to debate the consequences of automation in the workplace and on the job market. According to the Italian polling company SWG, 47% of Italians think automation will destroy more jobs than it will create, while 42% think the opposite. ⧉ **[IT 1]** Together with this, there is a growing awareness of the role of algorithms, their impact on how information is consumed and how they are used to create propaganda. At the international level, an increasing number of voices are calling for clearly articulated ethics principles to help govern Artificial Intelligence.

Italy has long suffered from a cultural and technological 'digital divide', and this gap affects the quality of debate on automation at all levels of society. In addition, civil society organisations are noticeable by their absence from the debate, and there is no specific entity dedicated to evaluate the impact of automation on policy.

Automated decision-making came to the fore and gained popular attention through the "Buona Scuola" education reforms in 2015. ⧉ **[IT 2]** These reforms included an algorithm for teachers' mobility that—as detailed in the section ADM in Action of this chapter—wrongly assigned teachers to new work locations in at least 5,000 instances, according to estimates by the Italian Labour Union. This caused uproar, resulting in several cases being brought before the Italian courts.

Algorithmic systems to detect tax evasion have also been a topic of discussion, however, these debates seldom depart from the 'Big Brother' narrative and are consistently tied to the risk of illiberal and excessive surveillance of taxpayers. This may be one of the cultural reasons why algorithms have not yet been widely adopted.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ White paper on "Artificial Intelligence at the service of citizens"

In March 2018, the so-called Italian Digital Transformation Team—led by former Amazon VP, Diego Piacentini, and made up of some 30 software developers and other technicians—were tasked with designing "the operating system of the country". They produced a "white paper on Artificial Intelligence at the service of citizens", which highlighted the potential of AI for public administration. ⧉ **[IT 3]** Among the recommendations in the white paper was a suggestion for the creation of a "National Competence Centre". This would, the report stated, constitute "a point of reference for the implementation of AI in the public administration that can provide predictions of impact and measurement of the social and economic effects of AI systems, in order to enhance the positive effects and reduce risks".

Trans-disciplinary Centre on AI

Among the other final recommendations included in the white paper, was a proposal to create a "Trans-disciplinary Centre on AI". This centre would be tasked with debating and

supporting "critical reflection on emerging ethical issues" associated with AI. It is also an issue which the Italian Chamber of Deputies is currently considering in a dedicated "Commissione d'Inchiesta" (parliamentary enquiry committee).

/ Expert Group for the Drafting of a National Strategy on Artificial Intelligence

The Italian Ministry of Economic Development (MISE) has issued a call for an "Expert Group for the Drafting of a National Strategy on Artificial Intelligence". The group will start operations in March 2019. Two of the group's goals specifically relate to automated decision-making. They are: "a comprehensive review of the legal framework with specific regard to safety and responsibility related to AI-based products and services", and an "analysis and evaluation of the socio-economic impact of development and widespread adoption of AI-based systems, along with proposals for tools to mitigate the encountered issues". ☑ **[IT 4]**

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ Nexa Center for Internet and Society

The Nexa Research Center for Internet and Society at the Politecnico Di Torino (Polytechnic University of Turin), is a research centre focusing on quantitative and interdisciplinary analysis of the impact the Internet has on society. ☑ **[IT 5]** In their project Data and Algorithm Ethics, Nexa researchers aim to "identify ethical issues raised by the collection and analysis of large amounts of data" and "focus on the problems arising from the increasing complexity and autonomy of some algorithms, especially in the case of machine learning applications; in this case the focus will be on unexpected and unintended consequences and on the responsibilities of designers and data scientists." ☑ **[IT 6]**

REGULATORY AND SELF-REGULATORY MEASURES

/ Declaration of Internet Rights

There is no legislation concerning automated decision-making in Italy. Guidelines and regulations are also missing for areas such as autonomous driving, algorithmic transparency and the use of automation in public policy and by law enforcement agencies.

However, general principles that might be applicable for ADM do exist within the Italian "Declaration of Internet Rights". This was devised by the "Commission on Internet Rights and Duties" in the Chamber of Deputies under the Renzi administration. ☑ **[IT 7]** The chamber was chaired by Laura Boldrini, who was then the Head of the Chamber of Deputies, and comprised of members of parliament, experts and journalists. The Commission published the Italian Internet Bill of Rights on July 28, 2015.

Of relevance to automated decision-making is Article 8. ☑ **[IT 8]** This explicitly states that "No act, judicial or administrative order or decision that could significantly impact the private sphere of individuals may be based solely on the automated processing of personal data undertaken in order to establish the profile or personality of the data subject". In ad-

dition, the need to avoid any kind of digitally-induced discrimination is repeatedly stated throughout the document, in terms of assuring net neutrality (Art. 4), "access to and processing of data" (Art. 5), "exercising civil and political freedoms" (Art. 10), access to platform services (Art. 12) and "the protection of the dignity of persons" (Art. 13).

ADM IN ACTION

/ Buona Scuola – algorithm for teachers' mobility

Within the framework of the "Buona Scuola" education reforms, an automated system was designed to evaluate how to manage the mobility of teachers for the 2016/17 academic year.

The system was developed by HP Italia and Finmeccanica and cost €444,000. ⧉ **[IT 9]** The algorithm was supposed to aggregate data on each candidate's score ("graduatoria") attributed by law depending on three specific criteria: work experience and roles, 15 preferred destinations, and the actual vacancies. The aim was to provide every teacher with the best possible outcome.

The system was supposed to prioritise the score. If teachers had a higher score, then their vacancy preference in a specific location would be considered before teachers with a lower score. Instead, the algorithm picked preferences without considering the scores. As more and more teachers found themselves assigned to destinations they didn't state in their preferences, many became disenchanted with the educational reforms. The resulting confusion provoked an uproar and made newspaper headlines.

Following the turmoil, a technical analysis ("Perizia tecnica preliminare") of the algorithm was performed by a team of engineers from the La Sapienza and Tor Vergata universities in Rome on June 4, 2017. ⧉ **[IT 10]** The report that followed detailed exactly why the mistakes happened—and possibly why they had to happen.

The authors stated that the algorithm was not up to the task of combining teachers and destinations in an automated way, because "even the most basic criteria of programming that notoriously apply" to the case at hand "were disattended". In fact, the code was so badly written that the engineers were left wondering how it was possible that the programmers had designed such a "pompous, redundant, and unmanageable system". The code was filled with bugs that could cause mistakes, such as the ones that happened.

The algorithm in question consisted of two parts and one of them was developed in the COBOL programming language. This is a language that was conceived in the early '60s and was mostly used in the business—not the government—sector. It was a choice that the authors of the technical analysis could not explain—unless one supposes a deliberate will to shield it from scrutiny, through the use of incomprehensible variable names and notes throughout the unusually long and chaotic code writing. This all resulted in several cases coming before the judicial courts throughout Italy, costing yet more of the state's time and taxpayers' money.

/ The "Abbiamo i numeri giusti" project

Designed by the Università Cattolica del Sacro Cuore in Milan, with the contribution of Merck, the "Abbiamo i numeri giusti" ("We have the right numbers") project is an automat-

ed system. It helps health institutions pick the most efficient and effective treatments for patients, while at the same time, it optimises public spending on data and evidence-based grounds. ☑ [IT 11]

The system, which aims to maximise patient engagement and medical compliance, will be tested and validated on a nationwide level, starting with trials in five regions (Emilia-Romagna, Veneto, Toscana, Lazio and Puglia) and using their existing databases. A high level 'Advisory Board', made up of institutional representatives at both national and regional levels, members of scientific societies and associations of patients, has been created to produce a report on the results of the experiment and to formulate ideas on how best to proceed. According to Merck, over 200,000 people die every year in Europe because they don't follow their treatment correctly. ☑ [IT 12]

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ RiskER – predict the risk of hospitalization

"RiskER" is a statistical procedure that combines over 500 demographic and health variables in order to 'predict' the risk of hospitalization. The automated system was developed by the Agenzia Sanitaria e Sociale (Health and Social Agency) of the Emilia-Romagna Region together with the Jefferson University of Philadelphia. ☑ [IT 13] It has been used on an experimental basis in 25 "Case della Salute" (public offices where administrative, social and healthcare services are provided at a local level), involving some 16,000 citizens and 221 general practitioners. ☑ [IT 14] The aim is to change the hospitalisation process of patients who, according to the algorithm, show higher health risks. ☑ [IT 15] It is hoped that the system will eventually be used in all 81 "Case della Salute" in the region.

The validation process has shown that the algorithm, which was revised in 2016, was good at predicting health risks in the 14 years and over age group and especially in older people, where the risks are higher. During the experiment, the algorithm grouped the population according to four risk categories, allowing doctors to identify high risk patients, and to contact, advise and/or treat them before their conditions became critical. Socio-economic variables might also be added to the algorithm in the future to increase predictive power. ☑ [IT 16] The "RiskER" algorithm is part of the EU's "Sunfrail" program which is aimed at helping the elderly.

/ S.A.R.I. – facial recognition system

The "S.A.R.I." (Sistema Automatico di Riconoscimento Immagini), or Automated Image Recognition System, is an automated decision-making tool used by the "Polizia di Stato", the national police force. ☑ [IT 17] This facial recognition technology has two functions: 1) ENTERPRISE, 2) REAL-TIME.

Through the use of "one or more" facial recognition algorithms, the ENTERPRISE function can match the face of an unidentified suspect in an image with facial images stored in databases administered by the police (which are "in the order of millions", according to the "Capitolato Tecnico" for S.A.R.I., the technical document in which the Interior Ministry detailed the system's desired requirements ☑ [IT 18]). It also provides a 'rank' of similarities among several faces with an associated algorithmic score. The ranking, whose criteria are not transparent to the public, can be further refined by both demographic and descriptive variables, such as race, gender and height.

The REAL-TIME function is aimed at monitoring video feeds in real-time in a designated geographical area, and again it matches the recorded subjects to an unspecified 'watch list' that contains individuals numbering in the hundreds of thousands. Whenever a match is found, the system issues an alert for operators.

The first solution was developed by Parsec 3.26 SRL for €827,000 ⧉ **[IT 19]**; the second was developed by BT Italia S.p.a./Nergal Consulting S.r.l. for €534,000. ⧉ **[IT 20]**

The S.A.R.I. system made headlines in September 2018, when it was revealed that two burglars in Brescia had, for the first time, been identified and subsequently apprehended as a result of the algorithmic matching obtained by the system. This prompted further analyses by law enforcement agents, who later confirmed that the suspects were, in fact, the two men suggested by the automated system.

Questions arose as to how many individuals are included in the databases that feed into the algorithmic choices (the police itself hinted at 16 million people, with no further specification as to their composition) and why. ⧉ **[IT 21]** There were also questions about the reliability of the algorithmic matching (in terms of false positives and false negatives), the safeguards for matched individuals, the cybersecurity and privacy profiles of the system.

/ ISA – Fiscal Reliability Index

Algorithms have slowly made their way into fiscal legislation in Italy. Many instruments have been devised over the last decade, with the aim of automatically checking for tax evasion.

Among them is the search for discrepancies between someone's stated income and her/his actual spending patterns ("Redditometro") or between declared invoices and standard of living ("Spesometro"). ⧉ **[IT 22]** In addition, the forthcoming "Risparmiometro", will try to automatically recognise anomalies between stated income, savings and standard of living by matching information from several databases.

As none of these instruments made a difference in combatting tax evasion, a new fiscal reliability index ("ISA", Indice Sintetico di Affidabilità) was created, and it became law in 2016/2017. The ISA is an arithmetic average of a set of elementary indices of fiscal reliability—including coherence with previous fiscal declarations, consistent compliance and an analysis of past and current anomalies—summarised using a 0-10 score that embodies the fiscal 'virtue' of the analysed subject within a timeframe initially set at the past 8 years. ⧉ **[IT 23]**

The idea behind the ISA is to revolutionise the logic of fiscal compliance, shifting from a 'punitive' approach to one that is based on 'rewards' for those who have consistently shown fiscal reliability, according to the index. As a result, taxpayers with higher scores will enjoy a reward system that includes an exclusion from certain fiscal inspections, and simpler conditions for compliance.

/ KeyCrime – predictive policing algorithm

KeyCrime is predictive policing software based on an algorithm of criminal behaviour analysis. ⧉ **[IT 24]** It was developed over the last 15 years by former investigative police of-

ficer Mario Venturi and designed to automatically analyse criminal behaviour, help identify trends and suggest ways of thwarting future crimes.

Reportedly, the system can now sift through some 11,000 variables for each crime. These range from the obvious (time, location and appearance) to the less obvious (reconstructions of witnesses and suspects during subsequent interviews, and even the modus operandi of the criminal). ⤴ **[IT 25]**

Video feeds are included in the analysed data. The idea behind the software is to rationalise the use of the police force and automatically deploy officers exactly where they are needed. However, this is not done by providing predictions about the areas in which certain crimes are mostly likely to be committed, but by giving clear indications as to where sought-after suspects might be about to strike, and therefore making it easier to apprehend him or her.
⤴ **[IT 26]**

In addition, the analytic capabilities of the KeyCrime software are used to attribute a series of crimes to the same suspect. Data gained in actual usage shows that when the number of crimes committed by the same suspect exceed four, the algorithm is 9% more effective than traditional methods.

Over the years, KeyCrime has been successfully deployed in the city of Milan. According to police data divulged to the press in 2015, this led to a 23% reduction in pharmacy robberies, and a suspect identification rate of 70% in bank robberies. ⤴ **[IT 27]** ⤴ **[IT 28]**

/ eSecurity – predictive urban security in Trento

This "eSecurity" project labels itself as "the first experimental laboratory of predictive urban security". ⤴ **[IT 29]** It was co-funded by the European Commission and coordinated by the eCrime research group of the Faculty of Law at the University of Trento, in partnership with the ICT Centre of Fondazione Bruno Kessler, the Trento Police Department and the Municipality of Trento. ⤴ **[IT 30]**

The project builds on the 'hot spots' philosophy. This is the idea that "in any urban environment, crime and deviance concentrate in some areas (streets, squares, etc.) and that past victimization predicts future victimization". It is an "information system for police forces and local administrations", modelled upon the predictive policing models adopted in the UK and the U.S., and aimed at providing complex automated assistance to law enforcement agencies and decision-makers.

The expanded "predictive urban security" model adopted by eSecurity employs algorithms that can: 1) use "past geo-referenced crimes" and smart city data, 2) consider "the concentration of victimisation, insecurity and urban disorder at city level", 3) try "to not only predict 'where' and 'when' specific types of criminality and deviance will occur but also to understand 'why' they will occur", and 4) are explicitly aimed at providing data-based insights to policy-makers for 'smart' and effective urban security planning.

Fabio Chiusi

is tech-policy advisor at the Italian Chamber of Deputies and Adjunct Professor in "Journalism and New Media" at the University of San Marino. As a reporter and columnist, he wrote about digital politics and culture for La Repubblica, L'Espresso, La Lettura del Corriere della Sera, Wired and Valigia Blu. He coordinated the PuntoZero research project on transparency in social media campaigning and wrote reports about the cybersecurity of IoT and the use of personal data during the 2018 Italian election. He is a Fellow at the Nexa Center for Internet & Society of the Politecnico of Turin, and author of essays on digital democracy and mass surveillance. His latest book is "Io non sono qui. Visioni e inquietudini da un futuro presente" (November 2018, DeA Planeta).

A report by the Scientific Council for Government Policy concludes that the existing regulatory framework for ADM should be upgraded significantly in order to provide sufficient protection of fundamental rights and safeguards against erroneous use.

The Digital Government Agenda stresses that even semi-automated decisions should comply with the principles of Dutch administrative law. Many of the calls for action in the agenda focus on research into the consequences and ethics of ADM.

AMSTERDAM

DEN HAAG

CAPELLE AAN DEN IJSSEL

EINDHOVEN

Some Dutch municipalities use the SyRI "Risk Indication System". Based on certain risk indicators, the software allegedly detects an "increased risk of irregularities", i.e. whether someone is breaking the law.

In July 2018, the East Brabant District Court, together with Tilburg Law School, appointed a professor to the new special chair in Data Science in the Judiciary. This person will be involved in performing (small) experiments and pilots with AI and ADM at the District Court.

AROUND THE COUNTRY

Some Dutch municipalities use ADM to prevent and detect truancy and early school-leaving by using algorithms that help decide which students should be paid a visit by a school attendance officer.

THE NETHERLANDS

BY GIJS VAN TIL

In the Netherlands, the impact of automated decision-making (ADM) and Artificial Intelligence (AI) on individuals and society is predominantly discussed as part of the larger Dutch strategy on digitisation. As a prominent part of this strategy, the Dutch Digital Government Agenda plays an important role in setting out how digitisation of public administration at multiple administrative levels should progress and how ADM and AI could be involved in this process. Rather than policy, many of the steps taken by the government in this field focus on intensive research into regulatory approaches and how public values and human rights can be safeguarded and protected alongside these developments. At the same time, aspects of ADM and AI already play a role in some of the current proposals for regulatory measures, for example, in an act setting out the regulatory framework under the Digital Government Agenda. Civil society and academia, in turn, are concerned about the pitfalls and ethics of the growing use and importance of ADM and AI. Alongside existing actors, such as the Rathenau Institute, and Bits of Freedom, a Dutch Alliance for Artificial Intelligence was recently launched to enable multi-stakeholder discussion on the responsible development, deployment and use of AI.

Meanwhile, more and more cases of ADM are already in use in the Netherlands. Notable examples include a big data analysis system called System Risk Indication (SyRI), as well as several predictive policing initiatives run by the Dutch police. In addition, the private sector already offers a range of different ADM-based solutions, for example in the field of alternative dispute resolution and credit/risk scoring.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

Unlike many other countries, no work has been done in the Netherlands on a national agenda specifically focussed on automated decision-making or Artificial Intelligence. In politics, the role of algorithms and ADM in society is being treated as part of a larger discussion on digital transformation.

Set out below are the parts of the current political debate that focus specifically on the use of ADM and that could have the greatest impact on individuals.

/ Dutch Digitalisation Strategy

The Nederlandse Digitaliseringsstrategie: Nederland Digitaal (Dutch Digitalsation Strategy) ⤴ **[NL 1]** highlights the opportunities that the use of automated decision-making and Artificial Intelligence bring, in particular, in making all sorts of processes more efficient (the processes in question are not detailed in the report).

The strategy stresses the need to keep up with the enormous competition from other countries in the field, and therefore cooperation between the public and the private sector is encouraged. Such cooperation is, for example, given form in the Commit2Data program. ⤴ **[NL 2]** This is a multi-year, national research and innovation programme built on the basis of a public-private partnership. It aims to further develop the use of Big Data around

themes such as logistics, health and data handling. The sub-programme Verantwoorde Waardecreatie met Big Data (Responsible Value Creation with Big Data or VWData) focuses specifically on research into the responsible use of applied technical and societal Big Data solutions.

In the field of Artificial Intelligence, a similar innovation programme is envisaged with the aims of developing knowledge about the technologies behind AI, its application, and the human factor (ethics, behaviour and acceptance). Transparency and explainability of algorithms are set to be an important theme in the programme.

At the same time, the strategy warns about the risks involved in the use of ADM and AI, for example in terms of autonomy and equal treatment. An entire chapter of the strategy is therefore dedicated to the safeguarding of public values and human rights. The need for their inclusion in the development and use of data and algorithms is emphasised.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ Agenda Digitale Overheid – Digital Government Agenda

As part of the larger digitisation strategy, the Dutch government published its Digital Government Agenda: NL DIGIbeter in July 2018. [NL 3] In the agenda, the government sets out how the digitisation of public administration at multiple administrative levels should progress. It acknowledges the increasing use and importance of ADM in public decision-making processes, service provision and governance—for example, in the process of issuing permits. It also encourages experiments in this area by governmental institutions, both on a national and sub-national level.

As in the broader Dutch Digitalsation Strategy, the agenda highlights the importance of the protection of constitutional rights and public values in cases where ADM is applied. It also stresses that even semi-automated decisions should comply with the principles of Dutch administrative law. In this light, many of the calls for action in the agenda are focussed on research into the consequences and ethics of ADM in inter alia administrative service provision.

Some of the research as set out in the Digital Government Agenda:

In August 2018, a research report instigated by the Dutch government—and conducted by Utrecht University—about the relationship between algorithms and fundamental rights was sent to the Dutch parliament. [NL 4] The report identifies problems that the growing importance of algorithms brings with respect to fundamental rights. These include the right to privacy, equality rights, freedom rights and procedural rights. The government was supposed to react to the report during the summer of 2018, but, as of November 2018, they have still not done so.

In June 2018, the government asked the Wetenschappelijke Raad voor het Regeringsbeleid (Scientific Council for Government Policy, WRR)—an independent advisory body on government policy whose members include prominent social scientists, economists, and legal scholars—to conduct research into the impact of Artificial Intelligence on public values. [NL 5] The request mentions the growing impact of AI, both in the public and private sector, and stresses the need for cross-domain research into the impact on public values. Besides the opportunities, the request also mentions specific risks associated with the use of AI, for example, in terms of discrimination regarding vulnerable groups.

The WRR has already published several reports that touch on the societal consequences of the use of ADM. These include the 2016 publication, "Big Data and Security Policies: Serving Security, Protecting Freedom" ⧉ **[NL 6]**, about the use of Big Data by the police, the judiciary and the intelligence services. It concluded that the existing regulatory framework should be upgraded significantly in order to provide sufficient protection of fundamental rights and safeguards against erroneous use. Specifically, the report concludes that the use of risk profiles and (semi-)automated decision-making should be regulated more tightly.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

Lastly, the Digital Government Agenda mentions that the Wetenschappelijk onderzoeks- en documentatiecentrum (Research and Documentation Centre of the Ministry of Justice and Security or WODC) will carry out research on the regulation and legality of algorithms taking autonomous decisions. The research, planned for 2018, is supposed to focus on the societal consequences of ADM in the near future and any regulatory interventions that might be required. Results have not yet been published.

/ Dutch Council of State – Raad van State

On August 31, 2018, the Dutch Council of State—the highest advisory body on legislation and governance to both the government and parliament—published a critical assessment of the Dutch Digital Agenda and especially of its ADM provisions. ⧉ **[NL 7]** According to the Council of State, the growing importance of ADM carries with it the risk that citizens cannot check which rules are being applied. In addition, it is no longer possible to determine whether the rules actually do what they are intended to do. Furthermore, citizens risk being profiled and confronted with decisions based on information of which the source is unknown. The agenda warns of the detrimental effect upon aspects of the Dutch constitution, of issues relating to the rule of law in general, and to the consequences for the position and protection of individuals in particular. As an antidote, the council, among other things, proposed that an administrative decision must contain explanations of which ADM processes (algorithms) have been used and what data has been taken from other administrative bodies. In addition, the council proposes that a human check must be allowed if a citizen objects to an automated administrative decision. This would be done in order to strengthen the position of citizens in automated and follow-up decision-making.

A reaction from the government to the assessment was sent to parliament on November 2, 2018. ⧉ **[NL 8]** The government acknowledges the risks identified by the Council of State and mentions that the assessment is in line with actions and processes already initiated by the government.

/ The use of Artificial Intelligence in the judicial branch

While the Dutch judiciary has yet to implement ADM in any tangible form, steps are being taken to allow it in the near future. The Dutch Digitalisation Strategy has already asked for cooperation between the judicial branch and the scientific field to research ways in which Artificial Intelligence can be applied responsibly. On March 29, 2018, the Ministry of Justice and Security organised a round table discussion on the use of Artificial Intelligence in the legal field. ⧉ **[NL 9]** Among the participants were scientists as well as delegates from civil society organisations and industry. In July 2018, the East Brabant District Court, together with Tilburg Law School, appointed a professor to the new special chair in Data Science in the Judiciary. This person will be involved in performing (small) experiments and pilots with AI and ADM at the District Court. ⧉ **[NL 10]** The Minister of Legal Protection

promised to send a letter to parliament in the autumn of 2018 on the possible meaning of ADM and AI for the judicial branch. As of November 2018, this letter has not been received.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ Bits of Freedom

Bits of Freedom ⧉ [NL 11] is a digital rights organisation focussed mainly on privacy and freedom of communication online. It strives to influence legislation and self-regulation, and to empower citizens and users by advancing awareness, use, and development of freedom-enhancing technologies. The organisation is very active in public and political debates on the use of ADM in, amongst other things, predictive policing, profiling and credit scoring. For example, they took part in the round table discussion on the use of AI in the judicial branch. ⧉ [NL 12] More recently, Bits of Freedom ran several campaigns to raise awareness about the growing influence of ADM in modern day society. For example, they presented a Big Brother Award to the Dutch national police for their predictive policing initiatives. Furthermore, they co-initiated Heel Holland Transparant (All of Holland Transparent) ⧉ [NL 13], an initiative aimed at illustrating the amount of data gathered by private scoring companies. They did this by making public the personal details of some well-known Dutch people where all the information originated from public sources.

/ De Kafkabrigade

De Kafkabrigade (The Kafka Brigade) ⧉ [NL 14] tackle redundant and dysfunctional bureau-cracy that prevents people from accessing the services they need and that constrains and frustrates public service staff. The Brigade has published a book called De Digitale Kooi (The Digital Cage) that illustrates the unwanted consequences of the increasing use of, among other things, automated decision-making mechanisms in public service.

/ Dutch Alliance for Artificial Intelligence

On October 11 2018, the Dutch Alliance for Artificial Intelligence (ALLAI) ⧉ [NL 15] was unveiled to the public. The alliance was initiated by three Dutch members of the European High-Level Expert Group on AI: Catelijne Muller, Virginia Dignum and Aimee van Wijns-berghe. The aim of the alliance is to enable a multistakeholder discussion on the responsible development, deployment and use of AI. ALLAI focuses on six broad themes that cover AI's advantages as well as the risks and challenges it brings: high quality AI; ethics of AI; social impact of AI; education and skills; rights and accountability, and AI for good.

/ Platform Bescherming Burgerrechten

Platform Bescherming Burgerrechten (Platform for the Protection of Civil Rights) ⧉ [NL 16] is a civil rights NGO consisting of a network of organisations, groups and individuals who join each other in striving to better guarantee and strengthen civil rights in the Nether-lands. It focuses particularly on respect for privacy rights, physical integrity, digital au-tonomy and the right to personal control (and possession) of personal data. Among other things, Platform Bescherming Burgerrechten is involved in a coalition that initiated legal proceedings against the System Risk Indication (Systeem Risico Inventarisatie or SyRi—see ADM in Action).

/ Privacy First

Privacy First ☑ [NL 17] is a non-governmental organisation (NGO) committed to promoting and preserving the public's right to privacy. The foundation has several focus areas (e.g. financial, online and medical privacy), some of which centre on the growing impact of ADM on society, for example through profiling. Privacy First is also involved in the coalition that initiated legal proceedings against the System Risk Indication (Systeem Risico Inventarisatie or SyRI—see ADM in Action).

/ Rathenau Institute

The Rathenau Institute ☑ [NL 18] is an independent knowledge institute that in recent times has published several influential reports on automated decision-making, Big Data and AI. These reports are frequently mentioned in political debates, for example in a debate that preceded the passing of the Digital Government Act. In 2017, the institute published a report called "Opwaarderen. Het borgen van publieke waarden in de digitale samenleving" ("Urgent Upgrade. Protect public values in our digitised society") in which it concluded that digitisation challenges important public values and human rights such as privacy, equality, autonomy, and human dignity. The report warned that government, industry and society are not yet adequately equipped to deal with these challenges. Also in 2017, the Institute published a report on "Mensenrechten in het robottijdperk" ("Human rights in the robot age"). More recently, it published the "Doelgericht Digitaliseren" ("Decent digitisation") report in 2018. This report included a collection of blog posts in which experts set out their views on decent digitisation, for example in terms of how we can stay in charge of algorithms. On the basis of these reports, the institute formulated four virtues that can help to deal better with digital technology: personalisation, modesty (i.e. awareness of the limits of digital technology), transparency, and responsibility.

/ Scientific research

As set out above, the political debate on aspects of automation has already prompted quite a number of reports and research. In the scientific field, the need to research the impact of automated decision-making across different sectors of Dutch society is acknowledged as well. Such a call can, for example, be found in the Nationale Wetenschapsagenda (Dutch National Research Agenda). ☑ [NL 19] Furthermore, automated decision-making is highlighted in multiple programme lines of the Digital Society Research Agenda by the Association of Dutch Universities' (Vereniging van Universiteiten or VSNU). ☑ [NL 20]

Notable recent research includes a report ☑ [NL 21] that is part of a research collaboration between the Universities of Amsterdam, Tilburg, Radboud, Utrecht and Eindhoven (TU/e) on automated decision-making, and which forms part of the groups' research on fairness in automated decision-making. The report provides an overview of public knowledge, perceptions, hopes and concerns about the adoption of AI and ADM across different societal sectors in the Netherlands. Furthermore, a PhD thesis, on Automated Administrative Chain Decisions and Legal Protection, was recently defended at Tilburg University. ☑ [NL 22] The PhD candidate found that in most cases, administrative chain decisions regarding income tax filings or the granting of child benefits severely lacked transparency and contestability—thereby risking to infringe on fundamental rights as well as on administrative principles of good governance.

REGULATORY AND SELF-REGULATORY MEASURES

/ Digital Government Act

The proposal for the Wet Digitale Overheid (Digital Government Act) [NL 23] was submitted to parliament in June 2018. The act sets out the regulatory framework beneath the Digital Government Agenda, providing rules on, among other things, the power to impose certain (technical) standards in the electronic communication of the government; data and information security; the responsibility for the management of facilities and services within generic digital government infrastructure (GDI), and digital access to public services for citizens and businesses.

/ Data Processing Partnership Act

In September 2018, a public consultation was completed on a proposal for the Wet gegevensverwerking door samenwerkingsverbanden (Data Processing Partnership Act). [NL 24] The aim of the act is to provide a basis for public-private cooperation and to make collaboration between the two easier. This applies to the processing of data, specifically when this processing is used for surveillance or investigation purposes, for example to prevent crimes, or to detect welfare fraud.

/ Dutch Road Traffic Act 1994

An amendment of the Wegenverkeerswet 1994 (Dutch Road Traffic Act 1994) passed into law in September 2018. Upon obtaining a licence, this permits experiments with fully autonomous vehicles. [NL 25]

/ Code Good Digital Governance

As part of the Digital Government Agenda the presentation of a Code Goed Digitaal bestuur (Code Good Digital Governance) is envisaged for mid-2019. [NL 26] This code should aim to provide guidance and set out rules for the collection and use of data, and the use of new technologies, in the public space, for example in the context of smart city initiatives.

ADM IN ACTION

/ ADM in alternative dispute resolution

A few private initiatives are available that offer a form of alternative dispute resolution, or arbitration, through a completely digitised procedure. [NL 27] The cases are handled solely by using digital and algorithm-based processes—the defendant never sees an actual judge. Automated decision-making plays a significant role in these procedures, and these private courts, sometimes mockingly named Robo-judge, are increasingly being used by debt collection companies. Some health insurance companies also apply them by imposing their use on their clients by including arbitration clauses in their terms of use. Recently, this application of digital, privatised dispute resolution has sparked controversy, particularly regarding the lack of due process and the non-transparent nature of such initiatives. The Minister of Legal Protection has been asked questions about this in parliament. [NL 28] In response, the minister commented mostly positive about digital alternative dispute resolution initiatives and downplayed the risks and detrimental effects.

/ Credit / risk scoring

An increasing number of private companies offer credit scoring services. ↗ [NL 29] These services are used by a variety of clients—such as health care insurance providers, energy companies, internet service providers and phone companies—to rate the creditworthiness of citizens. The credit scoring companies gather information on a large scale from a variety of sources to give an automated indication as to whether a potential customer can be trusted. The client of the credit scoring service can subsequently use ADM to decide whether, for example, a potential customer can have insurance or a phone subscription. These private credit scoring companies exist alongside an official and independent financial registration office called Bureau Krediet Registratie (Central Credit Registration Office or BKR). In most cases, the amount of data these companies collect far exceeds the amount available at the BKR.

/ Journalistic reporting

ADM in The Netherlands has also found its way into journalism. ↗ [NL 30] Several news outlets have implemented, or are in the process of implementing, 'recommender systems'. These systems semi-automatically decide which articles are shown to each individual visitor or subscriber to a news website. Among these outlets are RTL Nieuws, Het Financieel Dagblad, NU.nl and the Dutch Broadcast Foundation (NOS). Most notable among these is a kiosk-like online platform called Blendle that enables users to read articles from multiple newspapers and magazines on a pay-per-view basis. It recently introduced a subscription model that provides subscribers with twenty tailored articles per day. Apart from a few articles that are hand-picked by editors, the selection of these articles is mainly algorithm-based and dependent on a variety of data points (e.g. what articles a user has previously clicked on).

/ Law enforcement initiatives – Predictive policing

At the moment, multiple initiatives are in operation in the Netherlands centred around the use of predictive, algorithm-based methods to anticipate and prevent crimes. Most notably the national Police has—after having run pilots in Amsterdam and The Hague—rolled out a programme called Criminaliteits Anticipatie Systeem (Crime Anticipation System or CAS), which they built themselves. ↗ [NL 31] ↗ [NL 32] The aim of this system is to predict where and when crimes, such as burglary and street robbery, will take place. The system does this by analysing a wide variety of data, such as historic crime rates, data from Statistics Netherlands, and information about recidivists (e.g. dates of birth and addresses), after which the likelihood of these crimes occurring is indicated in the form of a heat map.

Other examples of predictive policing initiatives and pilots are the development of *RTI-geweld*. This is a risk prediction instrument used to estimate the future risk of violence of all persons appearing in the police's incident registration system. In addition, *ProKID* is a method that was introduced in 2013. It is aimed at identifying the risk of recidivism among twelve year olds who have previously been suspected of a criminal offence by the police.

/ Municipality-level projects

In recent times, on the lower administrative levels (especially in municipalities), a broad range of data-driven or algorithm-based initiatives have seen the light of day. It goes beyond the stretch of this report to give a detailed overview of all developments at this point, but over recent years many municipalities have, for example, launched smart city initiatives.

These initiatives collect a broad range of data from a variety of sources and for a variety of reasons, such as improving safety in entertainment districts and crowd control, but also to regulate air quality and to solve mobility issues. An important development in this regard is the creation by a coalition of (larger) municipalities in collaboration with industry and scientists of the NL Smart City Strategie ⤢ [NL 33] in January 2017.

ADM is also used in some municipalities to prevent and detect truancy and early school-leaving. This is done by using algorithms that help decide which students will be paid a visit by a school attendance officer. Similar initiatives exist to detect child abuse and/or domestic violence.

Other than using System Risk Indication (see below), some municipalities have also developed their own initiatives that revolve around the use of algorithms to detect welfare fraud. These programmes take into account data such as dates of birth, family composition, paid premiums and benefits history, as well as data from the Tax and Customs Administration, Land Registry and the Netherlands Vehicle Authority. Municipalities thus hope to increase the chances of identifying people committing welfare fraud.

An overview of initiatives can be found in the 2018 report Datagedreven sturing bij gemeenten (Data driven steering in municipalities) ⤢ [NL 34], which was initiated by the Association of Netherlands Municipalities. The report urges municipalities to share knowledge, and encourages them to cooperate in the roll-out of new initiatives.

/ Social Welfare – SyRI

Systeem Risico Inventarisatie (System Risk Indication or SyRI) is a big data analysis system that runs under the auspices of the Ministry of Social Affairs and Employment. ⤢ [NL 35] It is used on request by any of the so-called 'cooperation associations'. These include state institutions such as the employee insurance provider (UWV), the tax office (De Belastingdienst), the social security bank (SVB) and the immigration authority (IND), as well as a few Dutch municipalities to detect wrongly collected social benefits and other abusive use of the social welfare state. The aim of the system is to combat and prevent the unlawful use or recourse to public money or social security institutions or other income-related state benefits. In order for SyRI to operate, data provided by a citizen (for example to file a tax return) is combined with data from a variety of other sources. An algorithm, involving a risk model with several unknown indicators, then determines whether a citizen should be flagged because of an increased risk of irregularities, or potential fraud.

In recent times, SyRI has increasingly attracted negative attention. The subject has led to questions to the Minister of Legal Protection by members of the Dutch House of Parliament. ⤢ [NL 36] More importantly, a group of civil rights initiatives which gathered under the name Bij Voorbaat Verdacht (Suspected in Advance) recently started legal proceedings against the use of the software. ⤢ [NL 37] A trial was set to start in the final months of 2018.

Gijs van Til LL.M.

is project researcher at the Institute for Information Law (IViR) of the University of Amsterdam. In this capacity, he is involved in several of the Institute's ongoing research projects, ranging from the future of copyright to the legal implications of the use of Artificial Intelligence and automated decision-making. During his studies in both Private and Information Law, Gijs gained profound knowledge of the Dutch and European legal system and developed his interest in the intersection of law and technology. A paper which he wrote in the course of his study about the relationship between copyright and search engines was published in a Dutch scientific journal. He wrote his master thesis about the proposed use of self-/coregulatory measures to tackle online disinformation. Besides his work at IViR, Gijs is as board member of Clinic, a law clinic providing free legal advice to individuals and start-ups on issues ranging from intellectual property to privacy.

POLAND

In January 2018, the Polish Ministry of Justice introduced a system to randomly allocate court cases to judges. The system is currently the most visible and most-discussed case of ADM used by the public administration.

The Panoptykon Foundation was the first to initiate a debate on algorithmic decision making in Poland. It analysed and then intervened in a case concerning the use of algorithms for profiling in unemployment.

GDANSK

WARSAW

WROCLAW

The Ministry of Labour and Social Policy introduced an ADM mechanism that profiles unemployed people and assigns them to categories that determine the type of assistance a person can obtain from local labour offices.

In Wroclaw, one of Poland's largest cities, a system malfunction caused the algorithm to incorrectly allocate children to nurseries, preschool and schools.

POLAND

BY ALEK TARKOWSKI

Policy debates on the issue of algorithms and automated decision-making (ADM) have only recently been initiated by the public administration, and they have focused on the related concept of Artificial Intelligence (AI). Before 2018, there were no signs of a policy debate on ADM and related issues. However, that changed in spring 2018, when public interest increased.

In April 2018, Poland was one of the Member State signatories of the Declaration of Cooperation on Artificial Intelligence, which was spearheaded by the European Commission. ⧉ **[PL 1]** In June 2018, the Polish government announced that work will begin on Poland's AI strategy. In November, the Ministry of Digital Affairs published a document titled "Proposition for an AI Strategy in Poland" that includes an action plan for 2018-2019.

ADM solutions are employed to a limited extent by the Polish business sector. According to a report published in 2018 by the Sobieski Institute on "Artificial Intelligence and IoT in Poland" ⧉ **[PL 2]**, large Polish IT companies do not yet have the capacity to deploy AI or ADM solutions. Greater capacity to do this can be observed among start-ups, micro, and small and medium-sized companies, especially in the FinTech sector. In most cases, these solutions are still in the early phases of implementation and business life cycle. Some of the first ADM projects are also being implemented in the public sector.

There are few visible signs of discussions that concern social, civic or ethical implications of these solutions. By framing the debate around the general term 'Artificial Intelligence', Polish stakeholders are avoiding a more specific discussion about the functioning of algorithms and their influence on decision-making. The debate about business solutions that employ ADM is currently focused on a growth paradigm, and explores the potential for further developing this sector of the economy. Similarly, a debate among academics working in the field of Artificial Intelligence focuses on attaining scientific goals, or obtaining public funding either at the Polish or the European level.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ Polish AI Strategy and Action Plan for 2018-2019

At a conference in June 2018, Jarosław Gowin, Deputy Prime Minister and Minister of Science and Higher Education, declared that Poland will create its own Artificial Intelligence strategy. ⧉ **[PL 3]**

In July 2018, the Ministry of Digital Affairs invited stakeholders to participate in shaping the Polish AI strategy. Between August and October, around 200 stakeholders participated in four working groups that dealt with issues of data availability, financing, ethics and regulation, and education in relation to AI. On November 10, 2018, the Ministry of Digital Affairs published a report titled *Założenia do strategii AI w Polsce. Plan działań Ministerstwa Cyfryzacji* (Proposition for an AI Strategy in Poland. Action Plan of the Ministry of Digital

Affairs) ⧉ **[PL 4]**. The document presents recommendations of strategic and operational goals for the four areas defined above. It also includes a short action plan proposed by the ministry for the years 2018-2019. It should be noted that the document is not an official strategy—the strategic recommendations have not been in any way endorsed or approved by the government. It remains to be seen if and when the Polish government will present its AI strategy.

The issue of ADM has been extensively addressed by the "Ethics and Law" working group—although the term is not used directly, as the document mainly uses the general concept of AI. The group defines a need to ensure that as AI solutions are being implemented, basic rights are effectively protected. In addition, knowledge about the social impact of AI should be obtained, ethical standards defined, and high-quality regulation adopted for areas related to the implementation of AI. Transparency and explainability of algorithms are listed as one of the key legal challenges concerning the protection of basic human rights in relation to AI.

The proposed 2018-2019 action plan does not include either the implementation of ADM by the public administration or its regulation. The only AI solution proposed in the action plan is a chatbot for the National Qualifications Registry. It is telling that the plan includes little mention of regulatory measures for ADM and AI. Regulation is seen only as a means for providing more effective public support for research, prototyping and implementation of AI in the economy.

The government has declared that work on Poland's AI strategy will continue in 2019.

/ Visegrad 4 countries' thoughts on Artificial Intelligence

In April 2018, the V4 Group (Czech Republic, Hungary, Poland and Slovakia) published a shared position, titled "Visegrad 4 countries' thoughts on the Artificial Intelligence and maximising its benefits ahead of the release of the European Commission's Communication on the topic". ⧉ **[PL 5]** The document is rather short and general in its nature. Most attention in the document is paid to the issue of the availability of data, its security, and trust in data sources. The document also declares the importance of "formulating open, executable and recognised international norms concerning the research, development and implementation of ethically designed systems and solutions based on Artificial Intelligence technology". The document lists the following priorities:

- Digitisation as a priority of the European Union beyond 2020
- A pan-European initiative on establishing a framework for opening up data for innovation
- Opening the debate about a funding mechanism
- Uniform regulatory sandboxes
- Support for the reform of public administration in decision-making through AI solutions
- consideration of cybersecurity and trust issues
- examination of the impact on the EU labour force

It should be pointed out that both the regulation of ADM and its use by the public administration are addressed in these priorities. The regulatory sandboxes are understood as "digital, virtual environments, in which interested parties from different sectors can experiment with data and algorithms". Thanks to these sandboxes, regulatory bodies can observe the development of algorithmic solutions and make informed decisions concern-

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

ing their regulation. The issue of the use of ADM by the public administration is not further developed, save for the mention of the regulatory impact assessment as a potential process where ADM could be used.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

The issue of algorithmic decision-making has been addressed by Polish civil society organisations for several years. In 2018, ADM became a significant topic of public debate due to on-going European copyright reform and the issue of algorithmic content filters. These organisations are also active in the public debate and consultations connected with Poland's new AI strategy.

/ Panoptykon Foundation

The Foundation Fundacja Panoptykon ⧉ [PL 6] was the first to initiate a debate about algorithmic decision-making in Poland. It analysed and then intervened in a case concerning the use of an algorithmic system for profiling the unemployed by the Polish Ministry of Labour and Social Policy (see ADM in Action). The Foundation also addresses the issue of algorithms within its broader anti-surveillance activism.

/ ePaństwo Foundation

In 2018, Fundacja ePaństwo ⧉ [PL 7] launched the "alGOVrithms" project, which aims to map the use of ADM by the public administration in Central and Eastern European states. ⧉ [PL 8] The foundation focuses on the use of algorithms by public administration bodies.

/ Centrum Cyfrowe Foundation

The Foundation Fundacja Centrum Cyfrowe ⧉ [PL 9] has been addressing the issue of ADM with regard to content filtering, as part of its advocacy work on the new European Directive on Copyright in the Digital Single Market. The issue, dubbed "ACTA 2" (in reference to the infamous ACTA treaty, which sparked mass protests in Poland), became one of the key digital policy issues in the Polish public debate this year, alongside GDPR implementation. In autumn, the Foundation launched a public campaign on the issue, called "Internet is for the people". ⧉ [PL 10]

/ Coalition for Polish Innovation and Digital Poland

By 2018, two major cross-sector coalitions, Koalicja na rzecz Polskich Innowacji (Coalition for Polish Innovation[1]) and Digital Poland[2], launched working groups and a programme that dealt with issues related to AI and algorithms. Both of these coalitions are foundations, established with the goal of networking stakeholders (mainly business, but also civil society or academic institutions) on issues related to the regulation of digital technologies. These groups were largely formed in response to the growing interest of the government in

---

1　Website of the coalition: https://koalicjadlainnowacji.pl/en/
2　Website of the foundation: https://www.digitalpoland.org/en/

Artificial Intelligence and related issues. However, at the time of writing this report, neither of the two coalitions have published any position papers or recommendations on the issue.

REGULATORY AND SELF-REGULATORY MEASURES

At the moment Poland lacks any regulation on a national level that directly concerns algorithms or algorithmic decision-making.[3] Although several important case studies of the use of ADM can be observed at different levels of government, there doesn't seem to be any attempt to regulate this issue or to define standards. Neither exist programmes to raise awareness about the issue or to encourage implementation of ADM solutions. The solutions that are employed are also relatively simple algorithmic processes.

ADM IN ACTION

/ Ministry of Justice – System of random Allocation of Cases

On 1 January 2018, the Polish Ministry of Justice introduced the "System of Random Allocation of Cases" (*System Losowego Przydziału Spraw*), a digital system that, on a once-per-day basis, assigns cases to judges across the country. The system is currently the most visible and most discussed case of ADM used by the public administration.

Pilots of the programme were initiated in three cities in 2017. The launch of the system aligned with controversial reforms of the judicial system in Poland, leading to intense public scrutiny. According to anecdotal evidence, the allocation of cases was not random, with judges receiving extremely varied allocations, often seen by them as unfair. In particular, the random character of the algorithm has been questioned.

In 2017, the Ministry refused a freedom of information request by NGOs ePaństwo and Watchdog Polska to disclose the source code and the details of the algorithm that powers the system. [PL 12]

The Ministry has provided explanations on how the system functions and described how the algorithm works, but it refuses to make the source code publicly available. According to the ministry, system administrators have limited permissions and cannot interfere with the randomised selection, the working of the programme is overseen by the Appellate Court in Gdańsk, and all operations made by users are registered. The Minister of Justice has declared that "the selection will be made solely by a machine, a computer system that is blind like Themis, and chooses without emotions, without views or biases, and in a manner fully free from possible accusations of corruption". [PL 13]

In mid-2018, the Ministry admitted that the system has faults and, in the case of some judges, assigns cases unequally. The ministry promised to introduce changes to the algorithm, yet its exact functioning remains unclear and controversial. [PL 14]

---

3   This opinion is shared by the authors of the report: Polityka Insight, "Iloraz sztucznej inteligencji. Potencjał AI w polskiej gospodarce", 2018 [PL 11]

/ School systems – allocation of children at schools

In Wrocław, one of Poland's largest cities, a system malfunction caused the algorithm that allocates children to nurseries to make incorrect selections. The malfunction led to a rare case of public scrutiny of the use of ADM for allocating learners to nurseries, preschools and schools. ☑ [PL 15]

In a typical scenario, a citywide allocation system is used to place children in preschools. This is done based on information provided by parents, which can include data on such factors as the number of children, single-parent households, food allergies of children, handicaps, and material situation. The functioning of such systems is controversial especially at preschool level, where the lack of a sufficient number of preschools and the lack of any obligation for a child to be in preschool leads to a shortage of available places. Parents therefore feel subjected to an arbitrary system that allocates their children in a non-transparent and possibly unfair manner.

There is no data available on the scale at which such systems are employed. One of these systems, Platforma Zarządzania Oświatą (Education Management Platform) created by Asseco Data Systems ☑ [PL 16], is used by 4,500 schools and preschools in 20 Polish cities. The system offers a range of functions and there is no data available on what percentage of them use ADM solutions for the allocation of learners. These cases of ADM do not seem to draw much public scrutiny. However, anecdotal evidence shows that the decisions made by these systems are of great importance to the parents of children in preschools and schools.

/ Canard Speed Camera System

The Centrum Automatycznego Nadzoru nad Ruchem Drogowym (Road Traffic Automation Supervision Centre, CANARD) ☑ [PL 17] is a nationwide fotoradar (speed camera) system. It is connected to an IT system that uses image analysis algorithms to read license plates before it automatically fines drivers who are speeding. The centre was created in 2011 within the General Road Transport Inspection (Główny Inspektorat Ruchu Drogowego), and the traffic control infrastructure that uses ADM was launched in 2015. The system consists of 400 stationary speed cameras, 29 mobile units, 29 road-based traffic measurement systems, and 20 devices that register vehicles crossing intersections at a red light. In August 2018, the "Puls Biznesu" economic journal disclosed that the CANARD system fails to fine drivers of electric cars. ☑ [PL 18] Apparently, incomplete data on electric cars in the CEPiK national vehicle registry led the algorithm to identify electric cars as belonging to a special category (reserved, for example, for traffic control vehicles) that were exempt from speeding fines.

/ Ministry of Labour and Social Policy – profiling of unemployed people

In May 2014, the Ministry of Labour and Social Policy introduced an ADM system that profiles unemployed people and assigns them to three categories that determine the type of assistance that a person can obtain from local labour offices. ☑ [PL 19] The profiling mechanism is part of "Syriusz Std"—a nationwide IT system created by the IT department of the Ministry—which collects data on people who register as unemployed in labour offices, and on employees and their activities. The decision is made based on data collected through a questionnaire used by an employee of the labour office to question the unemployed person.

24 questions are used to collect data on two criteria that are taken into account: "distance from the labour market" and "readiness to enter or return to the labour market".

Fundacja Panoptykon, and other NGOs critical of the system, believe that the questionnaire, and the system that makes decisions based on it, profiles individuals based on personal data. Once the system makes a decision based on the data, the employee of the labour office can change the profile selection before approving the decision and ending the process. According to official data, employees modify the system's selection in less than 1% of cases. Furthermore, the system has been criticised for a lack of transparency about the distribution of public services, lack of oversight, and the alleged arbitrary nature of decision-making due to the simplification of data obtained from interviews. In addition, the system does not give the subjects of this ADM process the means to obtain data or an explanation concerning the decision, and the labour offices have limited means of analysing and evaluating the ADM process. Initially, the Ministry shared general information about the functioning of the algorithm with the Panoptykon Foundation, which investigated the case and which later made this information public. In 2016, the Foundation obtained detailed information about the questionnaire and the scoring algorithm through a freedom of information request. ⤢ **[PL 20]** In the same year, the Polish Commissioner for Human Rights asked the Polish Constitutional Tribunal to determine whether the profiling system is in line with the Polish constitution. However, the issues raised by the commissioner did not concern the ADM component, instead he questioned the lack of a redress mechanism and the basis for collecting personal data. ⤢ **[PL 21]** The case was solved in 2018, when the Constitutional Tribunal decided that the system needs to be better secured in a legislative act. ⤢ **[PL 22]**

/ National Health Fund and fraud detection

In September 2018, the National Health Fund (NFZ) announced that it will implement algorithms to enable closer scrutiny of public health expenditure. ⤢ **[PL 23]** Currently, regional NFZ offices use two different IT systems that are not integrated, and thus do not enable full data analysis. Data integration is therefore the first step of the project announced by the fund. Using the data, algorithms will compare an individual patient's history of medical procedures with standardised scenarios, in order to discover anomalies—which are potentially caused by fraud at health institutions. The NFZ estimates that by employing algorithms it will shorten the time needed to analyse all contracted health institutions from 16 years to 5 years.

⤢ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

/ One2Tribe

One2Tribe ⤢ **[PL 24]** is a Polish company that started out as a games developer, and then pivoted into providing a motivational platform based on gamification methods and behavioural psychology. In recent years, the company has been developing a machine learning solution that optimises the motivational platform to most effectively improve employee behaviour by appropriately defining challenges and prizes. Recently, they have also begun implementing algorithms that take into account individual work stress as a factor. In all cases, ADM is being used to tailor the system to the individual traits and preferences of employees. Their new venture, one2tribe labs, will use the same platform to motivate patients undergoing medical therapy. ⤢ **[PL 25]**

/ Nethone

Nethone ⎘ **[PL 26]** is a Polish company that works on ADM solutions for detecting financial fraud and has developed what it calls a "Know Your User" (KYU) solution (based on the concept of "Know Your Customer", which is a cornerstone of most financial products). The company makes predictions based on user-website interaction, hardware specifications, and other data points provided by their business partners. Machine learning solutions developed by Nethone are used to verify the identity of individuals making online transactions, in an effort to identify fraudulent transactions. The company is currently active mainly in Latin American, UK and US markets. In 2018, Nethone received a grant from the Polish National Centre for Research and Development to develop an ADM solution for combatting account takeover (ATO) attacks on bank account holders. The solution could in principle replace current authentication and monitoring methods. In September 2018, Nethone partnered with the Polish Association of Lending Institutions (PZIP), which gathers together more than 50 Polish institutions providing non-bank loans. The association recommends the Nethone system to its members as an additional anti-fraud solution.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

Alek Tarkowski

Sociologist, copyright reform advocate and researcher of digital society. Co-founder and President of Centrum Cyfrowe Foundation, a think-and-do tank building a digital civic society in Poland. Co-founder of Communia, a European copyright advocacy association, and of Creative Commons Poland. New Europe Challenger 2016 and Leadership Academy of Poland alumnus in 2017. Member of the Steering Committe of the Internet Governance Forum Poland, Program Board of the School of Ideas at SWPS University of Social Sciences and Humanities and Commonwealth of Learning's Center for Connected Learning. Formerly member of the Polish Board of Digitisation, an advisory body to the Minister of Digitisation (2011-2016), and member of the Board of Strategic Advisors to the Prime Minister of Poland (2008-2011). Co-author of the strategic report "Poland 2030" and Poland's long-term strategy for growth. Obtained PhD in sociology from the Institute of Philosophy and Sociology, Polish Academy of Science.

SLOVENIA

The Ministry of Public Administration invited all stakeholders in the field of AI to help find use cases and outline future needs and challenges for both the public and private sector regarding the use of Artificial Intelligence or automated decision-making systems.

A system used by the Slovenian Police at borders automatically matches travellers to "other police data" such as criminal files. The Human Rights Ombudsman and the Information Commissioner stated that such a system is not constitutional and filed a formal complaint in 2017.

The Ministry of Finance's financial administration uses machine learning to detect tax-evasion schemes and tax fraud, and rank as "risky" citizens likely to become tax evaders.

BRNIK

LJUBLJANA

AROUND THE COUNTRY

METIS is an "intelligent system for early detection of learning problems in primary schools". It uses machine learning to search for specific learning patterns and help teachers find "problematic" pupils, build on a detection model based on "more than 30 million grades" and "some other data".

SLOVENIA

BY LENART J. KUČIĆ

Slovenia has not adopted any strategic document on a national level regarding the use of Artificial Intelligence (AI), algorithms or automated decision-making (ADM) systems. However, in April 2008, the Ministry of Public Administration invited ⬈ **[SI 1]** all stakeholders in the field of AI to help them find use cases and outline future needs / challenges for both the public and private sector. According to the ministry, they will ask the government to form a working group that will outline a national strategy in the first half of 2019.

Slovenia also signed the European "Digital Day" declaration to "encourage cooperation in artificial intelligence, blockchain, eHealth and innovation." ⬈ **[SI 2]** ⬈ **[SI 3]**

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ Digital transformation and eGovernment

The Ministry of Public Administration issued two publications presenting Slovenian e-government's services and its "reference vision" for digital transformation. The first publication *eGovernment in Slovenia* describes government databases and services. ⬈ **[SI 4]** The second—Digital Transformation of Slovenia ⬈ **[SI 5]**—also mentions Artificial Intelligence and information infrastructure for future automated systems (e.g. driving, social security).

The Ministry for Public Administration is keeping most of the government's and citizens' data. According to a spokesperson, they have not yet tested or implemented any kind of autonomous systems but are using machine learning to improve their internal processes, e.g. the efficiency of government services and coordination of databases.

⬈ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ Financing Hate Speech

Slovenian online activist Domen Savič started a campaign documenting online ads from big Slovenian advertisers which appeared on various tabloid, political, astroturf, and disinformation websites. ⬈ **[SI 6]** He argued that advertisers are financing hate speech and called on companies to stop advertising on such websites. Most advertisers refused to comment. Some have tweaked their ad placement practices and silently removed their ads from problematic websites. One company admitted—or rather blamed—that "advertising programs" are responsible for ad placement and denied any responsibility for the placement of their ads.

In particular, Domen Savič focused on Telekom Slovenije—a majority state-owned national telecommunications company—because they were running ads on the website Nova24TV ⬈ **[SI 7]**, which was founded by the biggest Slovenian centre-right political party SDS. Savic noted that Telekom Slovenije presented itself as a responsible company promoting "respect

for all individuals". ⧉ **[SI 8]** Nonetheless, they decided to advertise on Nova24TV which regularly featured problematic, hateful, and false articles on migrants, ethnic and religious minorities (e.g. Muslims).

Telekom Slovenije stated that they do not promote any kind of intolerance with their advertising. Their only goal is to present their services to as many potential customers as possible. They also claimed that they do not want to interfere with editorial policy in any way.

The president of the Slovenian Advertising Chamber ⧉ **[SI 9]** explained that advertisers have to balance their corporate responsibilities and brand values on the one hand with editorial independence on the other. However, the Slovenian Advertising Chamber supports the idea of a "Code of Practice on Disinformation"—an initiative of European advertisers ⧉ **[SI 10]** proposed to the European Commission in September 2018. The Code of Practice calls for more transparent media buying practices and recognises the possibility that advertisers may involuntarily finance disinformation websites.

The two co-founders of the Slovenian content marketing company Zemanta ⧉ **[SI 11]** (which was recently taken over by the US company Outbrain) explained the functioning of the advertising "black box" and advertising algorithms. They admitted that advertising algorithms can sometimes be gamed and abused by disinformation sites. However, they did not agree with the argument that autonomous advertising mechanisms can be held responsible for financing hate speech. Indeed, the advertiser may be surprised by a certain "unfortunate ad placement". But they argue that it is relatively easy to fix this problem if an advertiser decides to better control its online ads. In short: "they (the advertisers) should learn to better use their tools".

Slovenian Prime Minister Marjan Šarec also joined the debate in November 2018. He published a public statement ⧉ **[SI 12]** and called for the advertisers to reconsider their marketing strategies when their clients promote hate speech (e.g. racism and homophobia). His appeal evoked radically polarised opinions. Political opponents, some advertisers, and right-wing and conservative media commentators accused him of censorship and political pressure on private companies. Left-wing and liberal media critics, on the other hand, embraced the PM's position that advertisers should not legitimise and finance hate speech.

⧉ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/automating-society*

/ Petition against automated weapons

In October 2018, a group of scientists, legal and technological experts organised a public debate on AI, autonomous weapons, and the risks of intelligent machines taking over the government. ⧉ **[SI 13]** They produced a joint statement and petition against research and the use of autonomous weapons. Such weapons should be banned in a way similar to chemical weapons, cluster bombs, and laser systems for blinding human combatants. The group also appealed to the Slovenian government to join the UN coalition against such weapons and support the recommendation of the resolution of autonomous weapons proposed by the European Parliament. ⧉ **[SI 14]**

/ #SOCRATECH – responsible IT development

In September and October 2018, the Slovenian online activist Domen Savič (he is a co-founder of an NGO called Državljan D—"Citizen D") organised an international tour of panel discussions and public talks about ethical uses of IT. ⧉ **[SI 15]** The tour included Slovenia, Croatia, Serbia, and Bosnia.

Speakers discussed the following topics:

- What is the meaning of ethical IT development?
- Why are smart cities dangerous?
- Why do we need to talk about Big Data?
- How can we protect ourselves against aggressive Artificial Intelligence and smart algorithms?

REGULATORY AND SELF-REGULATORY MEASURES

/ SI ST

The technical secretary at the Slovenian Institute for Standardisation (SI ST) ⤢ **[SI 16]** said that they have not yet received an application for standardisation of AI. However, they have registered the following standards in the field of autonomous systems / algorithms:

- SI ST ISO 9564-2:1995—Banking; managing and protection of personal identification numbers (PIN)
- SI ST-TP ETSI /SR 002 176 V1.1.1.2005—Electronic signature and infrastructure (ESI)
- SI ST ISO 10126-2:1995—Banking; encryption algorithms (DEA)
- SI ST EN ISO 12813:2016—Electronic collection of fees / communication for verification of compliance of autonomous systems (ISO 12813:2015)
- SI ST-TS CEN/TS 16702-1:2015, SI ST-TS CEN/TS 16702-2:2015—Electronic collection of fees / secure monitoring of autonomous tolling systems

ADM IN ACTION

/ Airport (border) police

On November 2, 2018, Slovenian daily Delo published a story about an Albanian citizen who tried to enter the Schengen area at Jože Pučnik international airport. ⤢ **[SI 17]** A "special algorithm" warned a border police officer to make additional checks on the suspect. The police found some falsified documents, Western Union money transfers and photographs for acquiring a false visa.

The Delo newspaper report stated that the police have acquired information about almost 800,000 airline passengers (so-called Passenger Name Records, PNR) since October 2017. The system tagged 8,928 passengers who were then thoroughly inspected before entering the country. The police stated that 40 per cent of those passengers were tagged as "not suspicious" and will not be reported next time they come to the border. Airline companies provided the data.

A police spokesperson explained they are not using algorithmic decision-making systems in this process. The system automatically matches a passenger to "other police data" such as criminal files. If a match is positive, a border police officer is informed and required to manually verify the information before inspecting a passenger. The police system also flags passengers who are using "unusual or illogical flights".

The Slovenian Human Rights Ombudsman and the Information Commissioner stated that such a system is not constitutional and filed a formal complaint in 2017. ⤢ **[SI 18]** ⤢ **[SI 19]** The Information Commissioner claimed that the adopted changes of the amended law on

the duties and powers of the police ☑ **[SI 20]**, which gave the police the power to gather the PNR, have legalised some excessive and inadmissible measures for gathering personal data without sufficient protection of citizens that have not been accused or suspected of any wrongdoings, e.g. terrorism or organised crime. They argued that all passengers should not be routinely scanned at the airport just because they are entering the state or landing during the transfer. The Human Rights Ombudsman supported their claims and the Slovenian Constitutional Court will therefore be required to rule on the constitutionality of the latest revision of the law on the duties and powers of the police.

/ Banking

The largest Slovenian bank NLB uses algorithmic systems in their operations to support decisions on granting loans and mortgages, to assist with fraud detection, and to improve the quality of their data (preventing wrong inputs and typos from clerks etc.). NLB's Chief Data Officer explained that their "quick-loan" offers (available via the mobile banking app) are almost entirely automated and only require final confirmation by a human. ☑ **[SI 21]** All Slovenian banks are required to access SI SBON—an information system for the exchange of data on individual debts before granting a loan. ☑ **[SI 22]** SI SBON is operated by the Bank of Slovenia and includes personal data and credit operation of individuals (credit data is held for four years). SI SBON has to be accessed by a person, not an automated system, according to NLB. Legislation thus limits the possibility for developing fully automated banking services for the time being. ☑ **[SI 23]**

NLB also said they do not discriminate against customers based on their demographics. "Slovenia is still a relatively egalitarian society and it does not make much sense to profile individual customers on their gender, place of residence, ethnicity or other categories," according to NLB's Data Officer.

/ METIS – detecting learning problems in primary and secondary schools

METIS is an "intelligent system for early detection of learning problems in primary schools." ☑ **[SI 24]** It is a joint project of the Jožef Stefan Institute and the Slovenian ministry of Education, Science, and Sport. ☑ **[SI 25]** The system has the potential to monitor pupils' grades (and absences) in 72% of all Slovenian primary and 90% of secondary schools. It will use machine learning to search for specific learning patterns and help teachers find "problematic" pupils. The researchers have built a detection model based on "more than 30 million grades" and "some other data". However, "teachers will keep all autonomy", according to spokespeople at the press conference. ☑ **[SI 26]** ☑ **[SI 27]**

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

METIS was first introduced in the summer of 2017. However, the representatives of the ministry of Education, Science, and Sport did not provide any additional information on when (or whether) the system will be fully implemented. According to a representative from the Jožef Stefan Institute, schools were required to obtain written consent from parents in order to run the system. This became the biggest obstacle for its adoption: "It was too complicated", the representative said.

/ Insurance

The biggest Slovenian insurance company Triglav ☑ **[SI 28]** uses algorithmic systems to help their insurance agents recommend their insurance products to customers, to detect fraud,

and to assess insurance risks. Nevertheless, they only use algorithmic tools as counsellors—they leave the final decision to their human agents, according to a spokesperson.

/ Tax-evasion and tax-fraud

The Ministry of Finance's financial administration ☒ **[SI 29]** confirmed that they are using machine learning to detect tax-evasion schemes and tax fraud, and to find errors in tax reports among other uses. They complement algorithmic models with real-life information from tax inspectors to better "teach" the algorithms. They are also developing mathematical tools for data mining and prediction analysis to find future improvements of the tax system (optimising, modifying, and collecting taxes).

A spokesperson from the financial administration confirmed they are ranking "risky" citizens who are more likely to become tax evaders. Their ranking depends on the type of a tax. A taxpayer who is likely to evade an income tax may not be the same person as someone who will try to evade a value added tax.

Lenart J. Kučić

is a journalist and podcaster at Pod črto: a Slovenian independent and non-profit media outlet focusing on investigative reporting, data journalism and in-depth stories. He also works as a researcher and an invited lecturer for several academic institutions including the Faculty of Social Sciences and Faculty of Arts (University of Ljubljana), Peace Institute, and the Jožef Stefan Institute. In 2008, he finished an MA programme at Goldsmiths College, London (with distinction) in Transnational Communication and Global Media. He was a staff writer (technology and media correspondent) at the biggest Slovenian national daily Delo for more than a decade. In 2015, he founded a podcasting network Marsowci where he co-hosts a book and a photography podcast. His recent work focuses on social and political impacts of new technologies, future of work, changes in media industry, and the politics of science. He lives and works in Ljubljana, Slovenia.
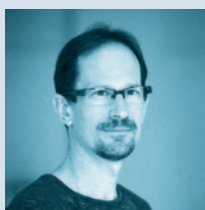
AROUND THE COUNTRY

The VeriPol tool is used to indicate the probability that a complaint made to the police is false by automatically analysing calls using natural language processing and machine learning techniques.

AROUND THE COUNTRY

The Spanish Public Employment Service (SEPE) uses an automated system to calculate unemployment benefits and to allocate job offers, interviews and training courses.

The system SAVRY is used in forensic criminology and was developed for assessing the risk of violence in adolescents (aged 12-18).

BARCELONA

The Fairness Measures Project is an international group of data scientists who develop fairness-aware algorithms and systems and provide relevant software and data sets to the research community.

SPAIN

BY KARMA PEIRÓ

Spain is still far away from having any regulation on automated decision-making (ADM). The Digital Strategy for a Smart Spain 2025 ⧉ **[SP 1]** is a guide that tries to show the way forward for technological innovation over the coming years. Other initiatives promoted by the Spanish government try to push private companies to make a commitment to techno-logical innovation, but their budget is small. Within the scientific community, there is an interest in advancing the ethical issues related to the application of ADM. Scientists want to ensure that a human, or team of humans, will always be able to make the final call, follow-ing any automated decisions. ⧉ **[SP 2]**

However, all applications of ADM are in a very experimental phase. Many of the examples collected here are announcements of predictions of what Artificial Intelligence might do in sectors such as banking, health or security, but not enough time has elapsed to obtain quali-tative results or to verify whether the ADM systems are as good as originally planned.

⧉ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ Digital Strategy for a Smart Spain 2025

The Ministry of Energy, Tourism and Digital Agenda is currently developing a Digital Strat-egy for a Smart Spain. This is based on the results of the current Digital Agenda for Spain (originally created in 2015) and addresses new rights. The document describes the applica-tion of Artificial Intelligence (AI) as a new opportunity: "The consolidation of platforms as fundamental agents of change and their role as arbitrators in the digital ecosystem will allow the advancement of industry, intensive automation, and the use of artificial intel-ligence".

The Digital Strategy for a Smart Spain 2025 ⧉ **[SP 3]** is based on five pillars:

- Data Economy (the ownership, value and ethics of data; development of digital tools that enhance the use of data in tourism, energy efficiency, diseases, problems of marginalisation, etc.)
- Ecosystems (public and private digital transformation; business and administrative digital transformation)
- Smart Regulation (sectoral economic regulation and defence of competition; revision and reform of taxation that combats the erosion of tax bases; avoidance or tax evasion that the digital environment makes possible; maintain protection of users' rights and consumer protection)
- Technological Infrastructure (extension of ultra-fast broadband connectivity; development of 5G networks and services)
- Citizenship and Digital Employment (improvement and alignment with the needs of the productive fabric of digital skills and competences and the promotion of STEM – Science, Technology, Engineering, Mathematics)

/ Activa Industria 4.0 Programme

The Activa Industria 4.0 programme ⧉ **[SP 4]** is supposed to assist the digital transformation of companies in Spain. Some 400 industrial companies from the 17 autonomous communities of Spain could benefit from this programme to advance their digital transformation and improve their competitiveness by adopting new enabling technologies. The programme is part of the Industry Connected Strategy 4.0 of the Ministry of Economy, Industry and Competitiveness.

With a budget of € 4 million, this initiative aims to understand the usefulness of applying technologies such as Big Data, web analytics, cybersecurity, cloud computing, robotics, sensorics, virtual reality and 3D printing.

/ National Plan for Scientific and Technical Research and Innovation

The main objective of this plan for 2017-2020 ⧉ **[SP 5]** is to identify and define strategic areas, strengths and contributions in the fields of research, development and innovation with the scientific and technical advice of experts and institutions. The plan gives financial support to research centres, universities, and companies that are adopting digital transformation projects, applied to processes like organisational innovation and societal challenges. The actions included in this National Plan contemplate the financing and co-financing by different administrations: national government and the European Structural and Investment Funds available for research and development and innovative activities. The research and development expenditure has been estimated by considering that total research and development spending reaches 2% of GDP in 2020, and the necessary convergence with the European average (EU-28).

/ The Artificial Intelligence lab of Aragon

The Technological Institute of Aragon (ITAINNOVA) ⧉ **[SP 6]** is a public institution funded by the Department of Industry and Innovation of the Government of Aragon. It will invest €3.5 million in the improvement of four laboratories before 2020. The goal is to start the improvement of the SHM (Structural Health Monitoring) laboratory: here the work is concentrated on intelligent systems, Artificial Intelligence and cognitive systems, and the Internet of Things (IoT).

/ Research and Innovation Programme on Advanced Digital Technologies

In November 2017, the Catalan government approved a research and innovation programme in advanced digital technologies. The aim of the programme was to promote technological development, establish synergies between research and innovation centres, improve the recruitment of talent, encourage investment, look into the impact of technology on administration and production, and to analyse how citizens lives are changing due to digitisation.

To obtain these results, the programme foresees the development of collaborative research, development and investigative projects in technologies such as 5G, the Internet of Things, Artificial Intelligence, computer vision, blockchain, and quantum technology. The programme is part of the SmartCAT strategy ⧉ **[SP 7]** of the Catalan government and the

research and innovation strategy for the smart specialisation of Catalonia (RIS3CAT). For 2018-2020, it has a budget of around €10 million.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION - CIVIL SOCIETY AND ACADEMIA

/ The Fairness Measures Project

 The growing use of automated decision-making has the potential to increase the risk of discrimination against disadvantaged groups. The Fairness Measures Project is a group of data scientists from Chile, Germany and Spain, led by Carlos Castillo (Pompeu Fabra University, Barcelona). The main goal of this group is to develop fairness-aware algorithms and systems ☑ [SP 8], and to provide relevant software and datasets to the research community through the website fairness-measures.org. The data sets cover several fields and applications such as finance, law and human resources, and provide common fairness definitions for machine learning. Up until now, the main results have included fairer algorithms for the ranking of people in searches within social networks and on job websites. This work has been received well in the media.

/ Barcelona Declaration for the Proper Development and Use of AI in Europe

This manifesto ☑ [SP 9]—lead by scientists Luc Steels (ICREA Research Professor) and Ramon López de Mántaras (Spanish National Research Council, CSIC)—is the result of the BDebate Conference, held in March 2017 in Barcelona. The manifesto proposes guidelines for the creation of a 'code of conduct' for AI practitioners, and it was very well received by the scientific community. The signatories of the manifesto believe that "AI can be a force for the good of society, but that there is also concern for inappropriate, premature or malicious use. This declaration is a step to ensure that AI is indeed used for the common good in safe, reliable, and accountable ways".

REGULATORY AND SELF-REGULATORY MEASURES

/ Science, Technology and Innovation Law

The Law 14/2011 ☑ [SP 10] regulates everything related to scientific and technological research. The regulation also proposed the creation of the Spanish Committee of Research Ethics—an independent and consultative body on professional ethics in scientific and technical research. The Committee issues reports, proposals and recommendations on matters related to professional ethics in scientific and technical research.

/ The Digital Strategy of the Catalan government

The Catalan government is preparing a regulatory framework ☑ [SP 11] on digital rights and duties. Although it is still at the draft stage, it focuses on collaboration in an open working group with institutions, professionals, experts and civil society. An announcement regarding the framework for the strategy was made in July 2018. It lists the rights and duties which the law will enforce, but it has not been defined yet. The aim is to present the strategy in 2019.

ADM IN ACTION

/ Automated weed mapping

In February 2018 a team of researchers from the Spanish National Research Council (CSIC) presented a system for early weed mapping. ⧉ **[SP 12]** It combines automatic learning techniques with photogrammetric techniques to help determine the height of the plants. "The algorithm generates maps of treatment that will help farmers in decision-making to improve crop management through the localised application of herbicides at the optimum phenological level, with substantial phytosanitary effects", said Ana Isabel de Castro Megías, a CSIC researcher at the Institute of Sustainable Agriculture in Córdoba. This study was carried out by a group of researchers led by Francisca López Granados from the Institute of Sustainable Agriculture in Córdoba, the Institute of Agrarian Sciences of Madrid, and the University of Salzburg (Austria).

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ Spanish Public Employment Service reduces benefits for the long-term unemployed

The Spanish Public Employment Service (SEPE) uses an automated system to calculate unemployment benefits and to allocate job offers, interviews and training courses. Over the last eight years, the number of people receiving benefits for long-term unemployment has decreased by more than 50 percent. However, it is unclear how much this automated system helped with this decrease. ⧉ **[SP 13]** One of the reasons for this is deficient data reconciliation. This is due to the simultaneous use of incompatible software ⧉ **[SP 14]**, meaning that a single mistake in a single field in a spreadsheet, in a database, or in the software running the allocation of the money can make a substantial difference.

/ VeriPol – a tool to detect false complaints to the police

An international team of researchers has developed the VeriPol tool which is used to indicate the probability that a complaint made to the police is false. It automatically analyses calls using natural language processing and machine learning techniques. The police claim that it is accurate 91% of the time. This success rate is fifteen percent higher than expert (human) police officers. The false negatives are 7.3% and the false positives 9.7%. ⧉ **[SP 15]** "This tool will help the police to focus research more effectively and to discourage false reports," says Federico Liberatore, a researcher at the Department of Statistics and Operations Research at the Complutense University of Madrid (UCM). ⧉ **[SP 16]**

VeriPol was developed for violent robberies, intimidation and theft. In recent years there has been an increase in the number of fabricated reports of these types of crime. From two sets of complaints, true in case one and false in the other, VeriPol automatically learns the central features of each set and trains a statistical model. ⧉ **[SP 17]** The application has been tested on over one thousand reports from 2015 provided by the Spanish National Police. It is the first time that a system like this has been implemented by the National Police. The project started in 2014 as a collaboration between UCM, the Carlos III University of Madrid, the University of Rome "La Sapienza", and the Ministry of Home Affairs of Spain.

/ Predictive evaluation by SAVRY

The goal of this investigation is to evaluate the predictive power and fairness of an expert assessment instrument called the Structured Assessment of Violence in Youth (SAVRY) and to compare it against standard machine learning (ML) algorithms. The system is used in forensic criminology and it was developed for assessing the risk of violence in adolescents (aged 12-18), but it was also seen to be effective in predicting the risk of general criminal recidivism. SAVRY plays a role in individual lives, and it influences the youth crime rate, as it can be used in intervention planning, such as clinical treatment plans or release and discharge decisions. Although these kinds of assessments do not intend to discriminate by gender or race, previous studies in the US—where similar systems have been used—have revealed unintended cases of discrimination.

The HUMAINT (HUmanity vs MAchine INTelligence) [SP 18] is an interdisciplinary research project that proposes evaluating fairness, taking into account the uncertainty of some predictions. In addition, it discusses the implications of different sources of bias for fairness and performance analysis. Researchers compare the performance of expert assessment with machine learning algorithms that also use information on defendant demographics and criminal history. "Our dataset comprises observations of 4,752 teenagers who committed offences between 2002 and 2010, and whose recidivist behaviour was recorded in 2013 and 2015. SAVRY is available for a subset of 855 defendants", says Carlos Castillo, a member of HUMAINT.

SAVRY is in general fair, while the machine learning models tend to discriminate against male defendants, foreigners, or people of specific national groups, says Castillo: "Machine learning could be incorporated into SAVRY, but if aspects of algorithmic justice are not taken into account, it could generate an unfair prediction." The evaluation [SP 19] showed that humans were much better than ADM, but that ADM can be more precise with the results.

/ RisCanvi – an actuarial risk assessment tool

RisCanvi is a statistical risk assessment system used in Catalan prisons, similar to LSI-R (Canada), Compass (US) and OaSys (UK). Although the tool makes predictions, the actual decisions are made by professional experts who sign off on any measures. The Department of Justice of the Catalan government launched the tool in 2010 and applied it to all inmates in all prisons, and not just for cases involving violent crime.

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

The Catalan government published a very positive evaluation of the tool's predictive ability. [SP 20] [SP 21] However, researchers, including Lucía Martínez Garay, urged caution. [SP 22]

/ Detection of hate in social media

Juan Carlos Pereira Kohatsu, a 24 year-old data scientist, created a tool to automatically detect hate on Twitter as part of his master's thesis. [SP 23] The tool was developed with help from the National Bureau for the Fight against Hate Crimes, part of the Ministry of the Interior, which is considering using it operationally to react to local outbursts of hatred.

/ Bismart – predictive algorithms to provide social aid

Bismart, a Spanish company, uses advanced data analysis systems to predict when it's necessary to provide aid to elderly people ⧉ [SP 24], without having to ask for it and before an emergency occurs. Smart Social Home Care is designed to go from a palliative to a proactive approach, and to distribute resources more efficiently. According to the company, the system aggregates data about social services, health, population, economic activity, utility usage, waste management, and more. Then, it uses this data to identify and predict groups and areas that will need urgent help. According to their website, the system is being used in Bilbao and Barcelona.

Bismart also provides services to detect illegal short-term rentals ⧉ [SP 25] (e.g. Airbnb), as well as a predictive policing solution. ⧉ [SP 26]

/ Jurimetria – a statistical and predictive tool for the legal sector

Jurimetria is a piece of statistical and predictive jurisprudential software that helps legal professionals analyse their cases. It systemises and extracts content from more than 10 million judicial decisions, coming from all instances and jurisdictional orders of Spain. ⧉ [SP 27] According to the company ⧉ [SP 28], half a million new resolutions are incorporated every year. In the same way, all the parameters of the judicial statistics of courts and tribunals of Spain are processed, updated, enriched and integrated continuously, including information on the duration, congestion, resolution, pendency and litigation in the legal system. The application allegedly allows users to extract and reveal unpublished procedural success patterns, starting from a complex framework of millions of jurisprudential documents to provide a quick and accurate response to all questions that may arise around a judicial process.

/ Machine learning to avoid financial fraud

BBVA, the second-largest Spanish bank, uses the services of Brighterion (a Mastercard company) to automatically detect fraud. ⧉ [SP 29]

/ Diagnosis of bipolar disorder

Researchers from the CIBERSAM (National Institute of Mental Health) used a self-learning algorithm to automatically detect bipolar disorder based on neuroimaging data. ⧉ [SP 30] The study, conducted on 3,000 patients in Barcelona, showed a level of accuracy of 65% using the algorithm. According to the authors, this was more accurate than diagnosis by a human, but too poor to be used in practice.

/ Mediktor – automated diagnosis

Mediktor is a tool for automated diagnosis. It relies on IBM technology but the Artificial Intelligence component is different from IBM's Watson. In 2017, Mediktor was tested on 1,500 patients at two hospitals in Barcelona and Madrid ⧉ [SP 31] and showed a success rate of 91.3%, in comparison with diagnosis by a medical professional. ⧉ [SP 32]

/ SavanaMed – records processing

SavanaMed is an Artificial Intelligence system that transforms the free text of clinical records into structured data. ⬀ **[SP 33]** It is already in use in Madrid, Castilla-La Mancha, Castilla y León, Valencia, Andalusia and Catalonia, which together make up two thirds of Spain's population. The system has processed more than 150 million clinical records.

/ Detection of diabetic retinopathy

A hospital in Barcelona is in the process of applying ADM to diagnose diabetic retinopathy—an illness that causes blindness—based on photographs of the retina. ⬀ **[SP 34]** The process has already been used elsewhere, but this experiment will be the first on a population from Southern Europe.

Karma Peiró

is a journalist specialized in Information and Communication Technologies (ICT) since 1995. Lecturer in seminars and debates related to data journalism, transparency of information, privacy, open dataand digital communication. She's co-director of the Diploma in Data Journalism at the Faculty of Communication and International Relations Blanquerna / URL (Barcelona). Also she's member of the Advisory Council for the Transparency of Barcelona City Council, member of the board for the Barcelona Open Data Initiative and member of the Council for the Governance of Information and Archives.

A governmental committee
evaluates the legal framework
for introducing automated
vehicles into ordinary traffic as
part of a radical transformation
of the transport sector.

In Jönköping in central Sweden, the start-up Einride started the first commercial use of an all-electric, automated truck. It drives back and forth between warehouses, covering six miles a day, some of it on public roads, at speeds below 40 km/h.

STOCKHOLM

JÖNKÖPING

TRELLEBORG

AROUND THE COUNTRY

Unionen, a white-collar
union, has conducted opinion
polls among its members to
map current applications of
automation in the service
sector. Based on the poll, the
union will develop a strategy
for how to approach ADM and
automation more generally.

Trelleborg has automated parts
of its social benefits management.
New applications are automati-
cally checked and crosschecked
with other related databases. A
decision is automatically issued
by the system.

SWEDEN

BY ANNE KAUN AND JULIA VELKOVA

The Swedish government aims to place the country at the forefront of the technological development around Artificial Intelligence (AI) and automated decision-making (ADM). In order to accomplish this, it kicked off several strategic initiatives since 2017, including extensive government reports on self-driving cars and a commission for Artificial Intelligence. It also provided broad funding for the improvement of knowledge and skills on Artificial Intelligence for universities and university colleges. Besides the Swedish government, there are a number of additional stakeholders who are investing in knowledge production and the development of AI on a large scale. One initiative that should be mentioned here is the Wallenberg Autonomous Systems and Software Programme (WASP). It granted €100 million to two leading universities in Sweden to develop machine learning, AI, and the mathematical apparatus behind them over the next eight years. At the same time, automated decision-making is increasingly implemented in public service institutions, especially on the municipality level. However, regulation lags behind this development and the public is largely unaware of automated decision-making in the public sector and welfare institutions. The private sector already offers a range of different ADM-based solutions, ranging from office workflow optimisation, credit scoring and juridical support to self-driving lorries and the automatic detection of dyslexia among school children.

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ Swedish AI strategy

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

In February 2018, the Swedish parliament published a strategy for the national adoption of AI. ⬀ **[SE 1]** It is framed as complementing the already existing digitisation strategy. The AI strategy argues that Swedish competitiveness and welfare depend on the development of AI and digitisation. It suggests that AI needs to be implemented in a 'sustainable' way, which rests on designing applications that are "ethical, secure, reliable and transparent". The strategy identifies as a crucial prerequisite the need for more expertise to develop and use AI in different parts of society including companies, municipalities, regional administrations and governmental agencies. It is, however, debatable to what extent different groups in society are actually involved. At present, most of the funding and strategic development takes place in the universities and as support for business environments. The need for expertise is linked to the need for educational institutions to produce experts, particularly engineers. ⬀ **[SE 2]** Cybersecurity and collaboration with the defence sector are also outlined as important, as well as building EU-wide partnerships around AI. Engineers are singled out and given priority, even though the strategy admits that the field requires the involvement of other professions.

The strategy stresses that automated decision-making by governmental agencies poses moral dilemmas and concerns. The strategy argues that it is necessary to work on AI at a pan-EU level and desirable to achieve both a dialogue and a leading role in setting EU-wide standards, best practices and infrastructure for implementing AI. ⬀ **[SE 3]**

/ New governmental committee for coordinated and accelerated development of policies related to the fourth industrial revolution (2019-2020)

In August 2018, the Swedish government created a committee in charge of developing policies and accelerating automation across industries in Sweden. ⬀ [SE 4] The government justified the creation of the committee with the accelerated adoption of "the fourth industrial revolution", which it described as characterised by "constant connectivity, smaller and more powerful sensors, artificial intelligence and machine learning". A major task for the group is to suggest policy developments. The committee references the governmental report on autonomous vehicles as an example that articulates important challenges to legislation, i.e. defining rules for not-yet-realised technologies and their areas of application. The committee is expected to support the work of the government by identifying policy challenges, contributing to the reduction of legal uncertainties and to speed up policy development, in particular in the following areas: precision medicine, connected industries (this refers to 'smart' factories or industrial manufacturing that has been digitised), connected and automated vehicles, transport and systems. The committee is tasked with producing analyses of barriers for the adoption of "the fourth industrial revolution", such as existing regulatory frameworks, to map the need for adjusting existing regulatory frameworks, to continuously come up with suggestions for the government regarding policy developments, promote a dialogue between relevant governmental agencies and regional actors, educational institutions, the non-governmental sector, and business for efficient collaboration concerning policy-developments. It is, however, not stated how, and more precisely which of these actors will be involved. In addition, the committee should seek collaboration with international actors such as EU institutions, the OECD, the World Economic Forum and other countries and international institutions. A referee group with representatives from the state and governmental agencies, regional actors, the business sector and organisations with experience in policy development will be assigned to support the committee. The committee is supposed to produce a mid-term report for the years 2019 and 2020 and a final report is due on December 31, 2021. The committee started its work at the beginning of October 1, 2018. The Head of the committee is Jon Simonsson, ⬀ [SE 5] a former entrepreneur and CEO, whose prior work for the government included being head of the section for innovation within the Ministry of Enterprise and Innovation.

/ Governmental report on policy development related to autonomous driving vehicles

In March 2018, a governmental committee presented a 1,314-page report ⬀ [SE 6] that evaluates the legal framework for introducing automated vehicles into ordinary traffic and makes concrete policy suggestions to improve it. The task for the committee was to consider a fast introduction of vehicles with automated functions as part of a radical transformation of the transport sector. In the English summary of the report, the authors state its main outcomes as follows: "In the opinion of the committee, multi-stage development of regulations is required to deal with developments in the field of automated, electrified and digitised mobility so that this development can take place in a safe, sustainable manner. The committee's proposals are intended to commence adaptation of the regulations so that these do not impede the development of new solutions for enhanced attainment of transport policy targets. One difficulty regarding this work has involved developing a regulation for a phenomenon that is not yet available on the market, namely fully automated vehicles capable of replacing the driver. The committee has attempted to suggest solutions that provide enhanced opportunities for testing and introducing advanced automated functions in vehicles in the short term, as well as certain fully automated vehicles. However, these so-

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

lutions can primarily be used even when a broader introduction becomes possible." ☒ **[SE 7]** The suggested solutions include both changes in the current regulations (for example in the regulation of penalties in traffic and for drivers' licences), as well as proposals for completely new regulations and laws (for example a new law and regulation on self-driving cars).

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ addAI initiative – a policy initiative

The addAI initiative ☒ **[SE 8]** is a collaboration between experts in academia, government and companies to discuss and explore the impact of smart algorithms and AI on society. Members of the initiative have organised workshops and speak at public events such as the Swedish Internet Days (every year in November).

The initiative works on questions like:
- **Sociology:** What is the best way to interact with AI and how may it change relations between humans?
- **Law:** How much responsibility should AI have? AI and the rule of law
- **Business:** What does an AI strategy mean for an organisation or a country?

/ The Internet Foundation

The Internetstiftelsen (Internet Foundation in Sweden, IIS) ☒ **[SE 9]** is an independent and public organisation that works—in their own words—towards a positive development of the Internet. The foundation is based in Sweden and is responsible for the country code top level domain .se as well as .nu. Besides supporting technological innovation and research about the Internet, the foundation publishes guides and teaching materials to enhance digital public education as well as knowledge production. In relation to ADM, the foundation has produced several publications on digital philosophy and algorithms for the general public.

/ Swedish Trade Union Confederation

The Landsorganisationen i Sverige (Swedish Trade Union Confederation, LO) published a report ☒ **[SE 10]** about how work has changed over the past 20 years and concludes that it has been increasingly impoverished. Impoverished work refers to work that requires less expertise, shorter training time, work that involves less variation in work operations, but rather involves repetitive tasks, work that does not require a lot of learning and where the employee has less options to work creatively. The report's main conclusion is that most workplaces are characterised by digital taylorism where assignments are increasingly standardised. This, of course, has consequences if standardised assignments are increasingly automated. The report refers to a number of positive examples where automation has led to the upskilling of workers and larger variations in work assignments.

/ The Union

Unionen (The Union) is a white-collar union that has conducted opinion polls among its members to map current applications of automation in the service sector. Based on the poll the union will develop a strategy for how to approach ADM and automation more generally. ⧉ **[SE 11]** The main questions of the poll are related to the number of jobs that have disappeared because of automation and how companies are enhancing the competitiveness of their employees in that context.

/ Wallenberg Autonomous Systems and Software Program

A private fund and initiative called the Wallenberg Autonomous Systems and Software Program (WASP) ⧉ **[SE 12]** has granted €100 million ⧉ **[SE 13]** to two leading universities in Sweden to develop machine learning, AI, and the mathematical apparatus behind them in the next 8 years. 300 doctoral students will be trained, and an additional €300 million will be added to the main funding of the program.

⧉ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

REGULATORY AND SELF-REGULATORY MEASURES

/ Legislative changes of the Administrative Act

A legislative change that entered into force in July 2018 allows governmental agencies to implement automated decisions. ⧉ **[SE 14]** The law regulating the procedures of Swedish governmental institutions (Förvaltningslagen) was amended with a new clause under the headline "How can a decision be taken?" It states that "A decision can be taken by a decision-maker responsible alone, or together with others, or in an automated manner". According to the commentary about the law, an automated decision means a decision made by machines without a person in the public authority taking any active part in the decision-making process. The decision builds upon propositions made in 2014 by a former Swedish committee for e-governance. As the law might be interpreted as being inconsistent with article 22 of the GDPR, an explicit commentary says that the specifics of the law include sufficient protection of individuals and the right to appeal.

Even though this change was implemented, the law does not cover the legislative framework that governs municipalities. This framework explicitly prohibits automated decisions-systems can only generate suggestions for decisions, but a person needs to approve them and assume responsibility. Indeed, the law makes an explicit distinction between automated decision-making and automated decision support. Where and when to draw the line between both is a topic of debate in Sweden at the moment. Municipal officials say that a change in the law to allow automated decision-making is 'desirable'. The law requires that a detailed motivation for the reasoning that led to a decision should be attached. A Swedish expert in municipal law suggests ⧉ **[SE 15]** that the requirement for motivation sets limits on automation which takes decisions with little consideration of personal circumstances. However, he adds that future systems could be made to cater for such details. In the meantime, the current legislative framework ⧉ **[SE 16]** makes illegal solutions of automated welfare decisions, such as those implemented in the municipality of Trelleborg, and multiplied in other cities in Sweden (see the ADM in Action section below). The lack of such clarity makes the forced relocation of employees, who used to work in the processing of welfare decisions, to other parts of the municipal authority potentially problematic.

ADM IN ACTION

/ Airhelp – Automation of juridical support

Since November 2017, Airhelp employs an advanced algorithm called Lara to analyse compensation claims from airlines and evaluate in real-time how likely it is that a passenger will be compensated for interference to flight and travel plans. ⧉ **[SE 17]** The algorithm bases the decision on a number of variables including flight status, flight statistics and weather forecast. In that way, the algorithm has automated parts of the responsibility in terms of arguing for compensation claims and decision-making.

/ Credit scoring

Lendo.se ⧉ **[SE 18]** provides an automated calculation of individual risk with 25 different banks and loan-givers in Sweden based on their unique risk-assessment criteria before taking out a loan. It then issues an automatically generated document that states whether the applicant should be given a loan or not. The document issued by Lendo is valid for 30 days. Similar services are offered by UC.se ⧉ **[SE 19]** and Bisnode.se ⧉ **[SE 20]**, among others, who automate risk assessment, but also offer other services like automated marketing based on algorithmic audience profiling that takes digital footprints as an input. ⧉ **[SE 21]**

/ Einreid – Automated lorries

The Swedish startup Einreid ⧉ **[SE 22]** pioneers "intelligent movement" through a cloud-based, real-time computing truck. The truck decides on its driving route in real-time using "big data analysis" that "integrates customer data, traffic data, etc. to optimise delivery time, battery life, and energy consumption—making the journey from A to B as efficient as possible"—according to the company's website.

/ Hedvig – Automated home insurance

The start-up company Hedvig ⧉ **[SE 23]** ⧉ **[SE 24]** ⧉ **[SE 25]** has gotten a lot of publicity in the media for having automated the filing of insurance claims. The app uses a voice input and after that it automatically writes the claim, sends it, and the insurance company automatically processes it and disburses the payment. At the moment, only home insurance is automated in this way.

/ Lexplore – Automatically detecting dyslexia in children

Lexplore ⧉ **[SE 26]** is employing AI to read and analyse the eye movement of children reading from a screen to detect dyslexia. In Sweden, the service is directly sold to both schools and municipalities. Lexplore is currently being piloted in different states of the US.

/ Social benefits – The Trelleborg model

Since 2017, Trelleborg has automated parts of its decision-making when it comes to social benefits. ⧉ **[SE 27]** New applications are automatically checked and cross-checked with other related databases (e.g. the tax agency and unit for housing support). A decision is automatically issued by the system. The number of caseworkers has been reduced from 11 to 3 and the municipality argues that they have considerably reduced the number of people receiving social benefits. They have been heading a pilot project to export their automa-

tion model to 14 additional municipalities and have received several innovation prizes. However, applicants and citizens have not been explicitly informed about the automation process. During the implementation process in another municipality, more than half of the caseworkers left their jobs in protest. ⬀ **[SE 28]**

Anne Kaun

is Associate Professor at the Department for Media and Communication Studies at Södertörn University, Stockholm. Her research combines archival research with interviews and participant observation to better understand changes in how activists have used media technologies and how technologies shape activism in terms of temporality and space. In her most recent book, "Crisis and Critique", she explores the role of media in the shaping of social movements and resistance to capitalism. Furthermore, she is interested in different forms of digital and algorithmic activism and is studying the consequences of automation in public service institutions. She also explores prison media, tracing the media practices and media work of prisoners since the inception of the modern prison system. Anne serves as a chair of the Communication and Democracy section of ECREA and is vice-chair of the Activism, Communication and Social Justice Interest Group within ICA.

Foto: Anna Hartvig

Julia Velkova

is a digital media scholar and post-doctoral researcher at the Consumer Society Research Centre at the University of Helsinki. Her research lies at the crossing between infrastructure studies, science and technology studies and cultural studies of new and digital media. She currently works on a project which explores the waste economies behind the production of 'the cloud' with focus on the residual heat, temporalities and spaces that are created in the process of data centres being established in the Nordic countries. Other themes that she currently works with include algorithmic and metric cultures. She is also the Vice-Chair of the section "Media Industries and Cultural Production" of the European Communication and Research Education Association (ECREA). Her work has been published in journals such as New Media & Society, Big Data & Society, and International Journal of Cultural Studies, among others.

The Data Justice Lab at the University of Cardiff examines the relationship between what it calls 'data-fication' and social justice.

Private company Imosphere provides systems to help town halls and the National Health Service to automatically calculate so-called personal budgets for people who receive social welfare payments.

LOUGHBOROUGH

CARDIFF

LONDON

The All-Party Parliamentary Group on Artificial Intelligence (APPG AI) suggests establishing AI auditing mechanisms, to incentivise the creation of ethics boards inside organisations, and emphasises the need of industry-led international collaborations, such as a forum on AI global governance.

The Science and Technology Committee (House of Commons) recommends to publish and keep updated a list of where algorithms "with significant impacts" are being used in Central Government, along with projects aimed at introducing public service algorithms.

UNITED KINGDOM

BY TOM WILLS

A survey of procurement data carried out for this report shows automated decision-making (ADM) in use or being introduced in a wide range of applications across the UK public sector, from military command and control to the supervision of schoolchildren while they use the Internet.

Civil society and academia are playing an important scrutiny role in certain areas – this chapter looks at examples from social care and policing – although with no central register of the use of ADM, there is likely much more going on without oversight.

A parliamentary inquiry dedicated to the topic of ADM took place in 2018. It identified the key issues but was thin on specific policy recommendations.

Existing legislation has limited bearing on ADM, although reviews in some areas, such as law that would affect self-driving cars, are underway.

Several new government institutions are being established which include ADM in their terms of reference. They are aimed variously at support for Artificial Intelligence (AI) as a source of economic growth and the development of rules and oversight mechanisms. The government professes its ambition to be a world leader in the research and development of AI and its regulation.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

POLITICAL DEBATES ON ASPECTS OF AUTOMATION – GOVERNMENT AND PARLIAMENT

/ AI Sector Deal

Office for Artificial Intelligence

In April 2018, the government announced that it would establish an Office for Artificial Intelligence as part of the "AI Sector Deal". The latter is a package of policies including financial support and tax breaks for research and development, aimed at boosting the Artificial Intelligence industry in the UK. The Office will be responsible for overseeing the implementation of the UK's AI strategy.

Demis Hassabis, co-founder of DeepMind, an Artificial Intelligence company now owned by Google, was appointed as an advisor to the Office for AI in June 2018. The announcement of Hassabis' appointment attracted some critical press coverage, which raised the issue of whether it is appropriate for an employee of a private company (Google) to be advising the government. [UK 1]

The Office does not yet have a web page [UK 2] and is currently recruiting staff. [UK 3]

AI Council

Also announced as part of the AI Sector Deal, the AI Council will be a counterpart to the Office for AI, composed of "leading figures from industry and academia". [UK 4] In June

2018, entrepreneur Tabitha Goldstaub was announced as its chair and spokesperson. ⬈ **[UK 5]** The council is tasked to promote the growth of the AI sector.

Centre for Data Ethics & Innovation (CDEI)

The government is in the process of establishing this new advisory body. It is expected to lead the work of dealing with ethical issues raised by new data technology, agreeing best practice and identifying potential new regulations. The government says it wants to lead the debate on these issues, not just in the UK, but around the world.

In its public consultation document on the role and objectives for the centre, ADM is one of three areas identified by the government as an example of an ethical issue raised by the use of data and AI. ⬈ **[UK 6]** While recognising the potential benefits to society of ADM, it mentions discrimination against job applicants and inequities within the criminal justice system as examples of issues that may arise as a result of ADM.

One of the six proposed themes for the centre's work is transparency, which is described with reference to the ability to interpret or explain automated decisions.

The creation of the centre was announced in autumn 2017. In June 2018, the government named Roger Taylor, co-founder of the healthcare data company Dr. Foster, as its chair ⬈ **[UK 7]**, and launched a consultation on the Centre's remit and its priority areas of work. The consultation closed in September 2018, and a response from the government is now pending.

/ House of Commons

Algorithms in Decision-Making Inquiry

The Algorithms in Decision-Making Inquiry ⬈ **[UK 8]** was launched in September 2017 by the Science and Technology Committee. As a *Commons select committee* inquiry, it consisted of a series of investigative public hearings carried out by a cross-party group of MPs. Its terms of reference included:

- The extent of current and future use of ADM across both the public and private sectors
- The scope for bias and discrimination in ADM, and how this might be overcome
- Whether ADM can be made transparent and accountable
- How ADM might be regulated

The inquiry published a report of its findings in May 2018. It found that the trends for big data and machine learning had led to an increase in the use of ADM across many areas, arguing that algorithms tend to increase in effectiveness and value as more data is used and combined.

The report identified many problem areas, but was rarely specific in advocating solutions. Instead, it mostly called on existing or forthcoming regulatory bodies to carry out further research.

The report recommended:

- Algorithms that affect the public should generally be transparent.
- New tools for algorithm accountability should be considered, perhaps including codes of practice, audits, ethics boards, or certification of algorithm developers.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

- Britain's privacy regulator should be adequately funded.
- The publication of Data Protection Impact Assessments should be encouraged.
- A procurement model for algorithms should be developed.

It suggested that the following areas should be reviewed or evaluated further:

- The scope for people to challenge the results of ADM
- Whether new data protection laws are needed
- Oversight of ADM by regulators in specific sectors

One concrete recommendation was that the government should publish and keep updated a list of where algorithms "with significant impacts" are being used in Central Government, along with projects aimed at introducing public service algorithms.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

The report was not covered in the British press[1], suggesting there is currently little political momentum behind ADM as a national issue.

All-Party Parliamentary Group on Artificial Intelligence (APPG AI)

This informal group of MPs and peers was established in January 2017. It has attracted sponsorship from a consortium of firms that serve as the group's secretariat through a private company called the Big Innovation Centre. The firms are Accenture, Barclays, BP, CMS Cameron McKenna Nabarro Olswang, Deloitte, EDF Energy, Ernst and Young, KPMG, Microsoft and PricewaterhouseCoopers. **[UK 9]**

The APPG published the first annual summary of its findings in December 2017. **[UK 10]** One of seven 'focus areas' was accountability. Here the report suggested organisations should be made accountable for decisions made by the algorithms they use; that the Centre for Data Ethics and Innovation (CDEI) should establish AI auditing mechanisms; that ethics boards inside organisations should be incentivised; and that industry-led international collaborations were needed, such as a forum on AI global governance, which should lead the global debate.

/ House of Lords

Select Committee on Artificial Intelligence

The Select Committee on Artificial Intelligence was formed in June 2017. It reported in April 2018 and the government published its response in June 2018.

The report was titled *AI in the UK: ready, willing and able?* **[UK 11]**

It was relatively lukewarm towards the idea of algorithmic transparency, arguing that achieving full technical transparency is difficult and often not helpful. It accepted that there would be "particular safety-critical scenarios where technical transparency is imperative", such as healthcare, autonomous vehicles and weapons systems. It said regulators in the relevant sectors must have the power to enforce this.

---

1   Factiva cuttings search for 'algorithms AND decisions' in UK national press, 30 October 2018. https:// drive.google.com/file/d/1v7C2N7He0czqwn28O9386TbQZ9TYjROW/view?usp=sharing

However, the report drew a distinction between transparency and the explicability of algorithmic decisions. Here the recommendations were more wary of new ADM technology, with the committee stating its belief that "it is not acceptable to deploy any AI system which could have a substantial impact on an individual's life, unless it can generate a full and satisfactory explanation for the decisions it will take". In cases such as deep neural networks, this may mean that the system should not be deployed at all, the report suggests.

The report also expressed concern about the qualifications to the ADM safeguards in what was then the Data Protection Bill (now the Data Protection Act 2018), which mean the rules only apply to decisions 'solely' made by machines.

/ National Data Strategy

In June 2018, the government announced that it would produce a National Data Strategy "to unlock the power of data in the UK economy and government, while building public confidence in its use". ⧉ [UK 12] No further details have been announced at the time of writing, although other government initiatives described in the next section, such as the Centre for Data Ethics & Innovation, are laying some of the groundwork for the strategy.

POLITICAL DEBATE ON ASPECTS OF AUTOMATION – CIVIL SOCIETY AND ACADEMIA

/ Alan Turing Institute

The Alan Turing Institute was established as the UK's national interdisciplinary research institute for data science in 2015. Thirteen British universities are members. It has started a research project called "Developing an ethical framework for explaining algorithmic decision-making", in conjunction with the Information Commissioner's Office (ICO). ⧉ [UK 13]

/ Big Brother Watch

Big Brother Watch is an independent research and campaigning group, founded in 2009. Its mission is to expose and challenge threats to privacy, freedoms and civil liberties amid technological change in the UK. One of its main campaigns is FaceOff, concerning the use of automated facial recognition in policing. It has also responded to government consultations concerning ADM and AI. ⧉ [UK 14]

/ British Computer Society Specialist Group on Artificial Intelligence

The British Computer Society (BCS) is a professional organisation for information technology practitioners and academics, with official status as a chartered institute. It is committed to "making IT good for society". Within the BCS, the Specialist Group on Artificial Intelligence was founded in 1980. It organises an international conference on AI. ⧉ [UK 15]

/ Data Justice Lab

The Data Justice Lab is a research lab at Cardiff University's School of Journalism, Media and Culture. It seeks to examine the relationship between what it calls 'datafication'—the collection and processing of massive amounts of data for decision-making and governance

⧉ *LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

across more and more areas of social life—and social justice. Its major research project DATAJUSTICE is looking into this question at a European level. It also investigates citizen scoring, the regulation of data-driven online platforms, and big data, inter alia. ⧉ **[UK 16]**

/ Privacy International

Privacy International is a charity dedicated to challenging overreaching state and corporate surveillance in the interests of security and freedom. It scrutinises UK government policy ⧉ **[UK 17]** as part of its global remit. Artificial Intelligence is one of the charity's topics of interest, and it identifies ADM as a problem area. The organisation lobbies for strong data and privacy regulations. ⧉ **[UK 18]**

/ The Society for the Study of Artificial Intelligence and Simulation of Behaviour (AISB)

The AISB was founded in 1964 and counts academics and professionals among its members. It organises an annual convention and publishes a quarterly newsletter. ⧉ **[UK 19]**

REGULATORY AND SELF-REGULATORY MEASURES

/ Data Ethics Framework

The Data Ethics Framework ⧉ **[UK 20]** is guidance from the Department for Digital, Culture, Media & Sport (DCMS), a central government department. It was published in June 2018 and sets out "clear principles for how data should be used in the public sector". It replaces the earlier Data Science Ethical Framework, which was published in 2016.

The guidance specifically addresses the question of ADM. As guidance, it does not carry the legal weight of statute. ⧉ **[UK 21]**

At the centre of the framework are seven data ethics principles. They cover public benefit, legislation and codes of practice, proportionality, understanding data limitations, robust practices and skill sets, transparency/accountability, and responsible use of insights.

Principle six is of particular relevance to ADM. Under the heading 'Make your work transparent and be accountable', it encourages civil servants to publish their data and algorithms.

The framework contains more detailed guidance for each principle and a workbook that civil servants can use to record ethical decisions made about a particular data project.

When it comes to algorithms, the workbook suggests publishing their methodology, metadata about the model and/or the model itself, e.g. on Github, an open software repository.

When procuring a system involving an algorithm from a private sector supplier, a series of questions is set out for civil servants to ask. These cover a range of issues that can lead to bias or lack of explicability and transparency.

/ Data Protection Act 2018

The Data Protection Act 2018 (DPA) became law in May 2018. It transposes the GDPR but also goes further than the EU legislation in relation to automated decision-making.

The Act states that decisions that "significantly affect" an individual may not be "based solely on automated processing unless that decision is required or authorised by law".

When decisions are made solely through automated processing, the act stipulates that data controllers must notify data subjects in writing. ⧉ **[UK 22]** It also provides for a right of appeal against such decisions, under which a data subject can request that the decision be reconsidered or made anew without solely automated processing.

Privacy International has argued that the act contains "insufficient safeguards" in relation to ADM. ⧉ **[UK 23]**

/ Review of laws on automated vehicles

In March 2018, the government announced ⧉ **[UK 24]** a review of driving laws "to examine any legal obstacles to the widespread introduction of self-driving vehicles". The review will be conducted by the Law Commission of England and Wales ⧉ **[UK 25]** and the Scottish Law Commission ⧉ **[UK 26]** and is set to take three years.

The Law Commissions are state-funded bodies charged with ensuring that the law in general is "as fair, modern, simple and as cost-effective as possible". It can make recommendations that are then considered by parliament.

The commissions said they aim to publish consultation papers by the end of 2018, which will seek to identify key issues to be investigated. The first year of the project will also include an audit of the current law.

The review may lead to increased public debate around self-driving cars and the automated decisions they make. In particular, the Law Commissions say they will highlight decisions to be made around ethical issues. Assuming the commissions identify the need for new legislation, this is likely to push ethical questions around self-driving cars into the political arena.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

OVERSIGHT MECHANISMS

/ Information Commissioner's Office

The Information Commissioner's Office (ICO) is the UK's data protection regulator, funded by the government and directly answerable to parliament. It oversees and enforces the proper use of personal data by the private and public sectors.

The ICO website provides guidance for organisations on all aspects of data protection, including requirements deriving from the GDPR and Data Protection Act. Its pages on ADM explain the requirements of GDPR Article 22 and encourage businesses to go beyond this, by telling their customers about the use of ADM to make decisions that affect them. ⧉ **[UK 27]**

Under the Data Protection Act 2018, the ICO was given increased powers to issue fines to organisations that break the law. At the time of writing, the Information Commissioner's Office was still working on updating its advice for data controllers to reflect the content of the DPA 2018 including new provisions relating to ADM. ⧉ [UK 28]

ADM IN ACTION

/ Facial recognition used by police forces

The civil society organisation Big Brother Watch has researched the introduction of automatic facial recognition systems by police forces in the UK. ⧉ [UK 29]

These systems take images from CCTV cameras and scan the faces of passers-by to see if they appear on databases of individuals of interest to the police. In a typical application, when the system detects a match, police officers may apprehend the person for questioning, search or arrest.

Unless the match can be readily discounted, human police officers are likely to follow the decision and act on suspicions raised by the system.

After raising concerns over its discriminatory potential, Big Brother Watch was granted limited access to observe the operation of the system at the Notting Hill Carnival 2017. According to the NGO's report, the police said that in the course of a day around 35 people had been falsely matched by the system. In these cases, officers took no action because in reviewing the images it was obvious that the wrong person had been identified. However, "around five" people—who later turned out to have been identified by the system in error— were apprehended and asked to prove their identity.

Big Brother Watch reported that there was only one true positive match over the entire weekend—but even then there was a catch:

> "The person stopped was no longer wanted for arrest, as the police's data compiled for the event was outdated."

In this case a civil society organisation has been able to draw attention to the flawed use of ADM. But the figures that support their case are not routinely published. It was only through concerted lobbying and tenacity that they were able to obtain them. This shows the important role that civil society organisations can play in ADM accountability. However, in the absence of a central public register of ADM, no one can say how many other systems are being implemented without oversight.

/ Personalised budgets for social care

In the UK, people who need social care—practical support because of illness, disability or age—can approach their local town hall for help.

Town halls in England have started using automated decision-making systems to help determine how much money should be spent on each person, depending on their individual needs. The resulting figure is known as a personal budget.

The impact of these decisions on people's lives is enormous. There has been no discussion in the media about the specific role of ADM in personal budgets. But this BBC report from 2010 ⌷ [UK 30] illustrates what is at stake:

> Graeme Ellis, who is registered blind and is a wheelchair user, has been on a personal budget for more than a year and was originally assessed as needing £21,000. [...]
> But then after being reassessed, he got an email from his social worker telling him that his council would have to cut their contribution by £10,000.
> He told the BBC's You and Yours programme he was frightened he was going to end back in the position he was in four years ago.
> "I'm frightened about the effect that being housebound will have on my well-being because being able to get out of the house and do things is one of the things that enables me to carry on."

Since 2007, governments of both main political parties in the UK have encouraged town halls and the NHS to start using personal budgets. ⌷ [UK31] By 2014-15, around 500,000 people receiving social care through a town hall were subject to a personal budget. ⌷ [UK32]

It is not known exactly how many people have had their personal budgets decided with the help of ADM. However, one private company, Imosphere, provides systems to help decide personal budgets for many town halls and National Health Service (NHS) regions. ⌷ [UK33] Its website says that around forty town halls (local authorities) and fifteen NHS areas (Clinical Commissioning Groups, which also have the power to allocate personal budgets) across England currently use the system, and that it has allocated personal budgets worth a total of over 5.5 bn. ⌷ [UK34]

What is the impact of the ADM?

Research in the Journal of Social Welfare & Family Law ⌷ [UK34] found that automated personal budget decisions did not always correspond to people's needs; that they could be used as a mechanism for implementing spending cuts; and that the algorithmic nature of the system led to a lack of transparency.

The paper serves not only to illustrate how flawed ADM decisions can adversely impact people's lives, but also how ADM systems might be scrutinised and what obstacles are sure to arise in other domains of ADM accountability research.

In addition, it shows the importance of the social and political climate in which ADM systems are used. In this instance, it can be argued that an ADM system has served as a Trojan horse for spending cuts and the outsourcing of decision-making to the private sector. This might not have been so if the decision-making rules behind ADM were made transparent to the many charities and campaigning organisations that advocate for the rights of social care service users.

LINKS: You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society

/ Procurement in the UK public sector

For this report chapter, an analysis of procurement notices published in the Official Journal of the European Union (OJEU) was conducted.[2] This provides a snapshot of some of the areas of the UK public sector where ADM is in use, planned to be introduced or under research and development.

It is not intended to be a comprehensive or a representative sample, but it does give an indication of the wide scope of ADM in the UK in terms of the sectors and the types of problems to which it is applied. The notices may be solicitations to tender or relate to systems already contracted.

The table shows that automation is being applied to high-stakes decisions in a diverse range of areas including industrial control systems, healthcare and the safeguarding of children.

Table: Selected OJEU procurement notices for ADM systems from UK public bodies, 2018

LINKS: *You can find a list of all URLs in the report compiled online at: www.algorithmwatch.org/ automating-society*

Sector	Application	Procuring body
Finance	Sanctions list screening [UK 35]	Financial Services Comp. Scheme
Health	High Risk Human Papillomavirus testing [UK 36]	NHS Scotland
Health	Physiotherapy triage [UK 37]	NHS - West Suffolk CCG
Health	Ventilator control [UK 38]	NHS Scotland
Health	Symptom checker (Babylon) [UK 39]	NHS – Hammersmith & Fulham CCG
Law enforcement	Biometric matching [UK 40]	Home Office
Law enforcement	Detection of migrants in lorries [UK 41]	Home Office
Military	Command and control [UK 42]	Ministry of Defence
Schools	Alert unsafe computer activity of children [UK 43]	Education Scotland

---

2   The OJEU can be queried using a publicly accessible web interface at https://ted.europa.eu. The following query was entered into the Expert Search form. The Search scope was set to 'Archives'. This returned all notices posted in the UK including one of the following keywords: algorithm, artificial intelligence or machine learning. The descriptions of the product or service being procured were then manually reviewed and those that appeared likely or certain to involve ADM were selected.
    CY=[UK]
AND(FT=[algorithm*]
    OR FT= ["artificial intelligence"]
    OR FT= ["machine learning"])

Transport	Rail traffic management ↗ [UK 44]	Network Rail
Transport	Transport management ↗ [UK 45]	Cambridgeshire County Council
Utilities	Control of gas network ↗ [UK 46]	Northern Gas Networks
Utilities	Control of water network ↗ [UK 47]	South East Water
Utilities	Control of water network ↗ [UK 48]	Bristol Water

Source: OJEU, AlgorithmWatch research

Tom Wills

is a freelance data journalist and researcher based in Berlin. His speciality is using computational techniques for journalistic research and analysis—a type of data journalism. Previously he led the data journalism team at The Times of London, using computational techniques to drive investigations on topics ranging from Westminster politics to the gender pay gap. Using tools such as the Python coding language to design algorithms for the purposes of public interest journalism has given him an insight into the perils of automation. He has also reported on the consequences of data-driven decision making, including a major investigation into the World-Check bank screening database which resulted in stories published in six countries as part of an international collaboration. Tom graduated from City University, London with a master's degree in Investigative Journalism in 2012.

ABOUT US

TEAM

/ **Brigitte Alfter** / Research network coordinator

Brigitte Alfter is an award-winning Danish-German journalist specialising in European affairs. After a career in local and national Danish media, including a position as Brussels correspondent, she turned to cross-border collaborative journalism with ad hoc teams as well as with existing networks such as the ICIJ. She is the co-founder of several cross-border journalism projects and structures such as the Farmsubsidy.org project, the Scoop-project, Journalismfund.eu, the European Investigative Journalism Conference & Dataharvest, the Investigate Europe team and the European Journalism ARENA. Brigitte has published a book about cross-border collaborative journalism in Danish and German. An English version will be published in 2019.

/ **Beate Autering** / Layout

Beate Autering is a freelance graphic designer. She graduated in design and runs the beworx studio with Tiger Stangl. Together they create designs, graphics and illustrations and also provide image editing and post-production services. Their clients include iRights, mdsCreative, Agentur Sehstern, UNESCO World Heritage Germany and visitBerlin.

/ **Sarah Fischer** / Team Bertelsmann Stiftung

Sarah Fischer works on the "Ethics of Algorithms" project at the Bertelsmann Stiftung. The project deals with the social consequences of algorithmic decision-making and aims to contribute to the design of algorithmic systems that lead to more participation for all. Previously, she was a postdoctoral fellow in the postgraduate program "Trust and Communication in a Digitalized World" at the University of Münster. There she worked on transparency and trust in search engine algorithms. On the same programme, she earned her doctorate in communication science on the subject of trust in health information on the Internet.

Photo: Josh White

/ **Graham Holliday** / Copy editing

Graham is a freelance editor, media trainer and author. He has worked in a number of roles for the BBC over the past fifteen years and he was a correspondent for Reuters in Rwanda. He works as an editor of non-fiction books, reports, documents and promotional literature for a variety of companies, NGOs and authors. He also works as an editor for CNN's Parts Unknown and for Roads & Kingdoms—the international journal of foreign correspondence. The late Anthony Bourdain published Graham's first two books which were reviewed in the New York Times, Los Angeles Times, Wall Street Journal, Publisher's Weekly, Library Journal and on NPR among other outlets.

Photo: Marion Kotlarski

/ **Nicolas Kayser-Bril** / Additional editing

Nicolas Kayser-Bril, 32, is a Berlin-based, French-German data-driven journalist. He created the data journalism team at OWNI, a French news startup, in 2010, then co-founded and led the data journalism agency Journalism++ from 2011 to 2017. He now splits his time between teaching programming at HMKW (Berlin), helping fellow journalists in their data-driven investigations, consulting for various organizations and writing history books.

Photo: Jan Voth

/ **Ralph Müller-Eiselt** / Team Bertelsmann Stiftung

Ralph Müller-Eiselt is a Director at the Bertelsmann Stiftung. He holds a B.A. in International Relations from Dresden University and graduated with an M.P.P. from the Berlin-based Hertie School of Governance in 2010. Before joining the Bertelsmann Stiftung, he was involved in several large-scale change projects in the public sector. For several years, he has been heading both projects in the field of digital education and the foundation's taskforce on policy challenges and opportunities in a digitalised world. His latest project "Ethics of Algorithms" takes a close look at the consequences of algorithmic decision-making and Artificial Intelligence in today's society.

/ **Kristina Penner** / Additional editing

Kristina Penner is the Executive Advisor at AlgorithmWatch. Her research interests include ADM in social welfare systems, social scoring and the societal impacts of ADM as well as the development of participatory and empowering concepts. Her analysis of the EU border management system builds on her previous experience in research and counselling on the implementation of European and German asylum law. Further experience includes projects on the use of media in civil society and conflict sensitive journalism as well as stakeholder involvement in peace processes in the Philippines. She holds a master's degree in International Studies / Peace and Conflict Research from Goethe University in Frankfurt.

Photo: Manuel Kinzer

/ **Matthias Spielkamp** / Editor

Matthias Spielkamp is co-founder and executive director of AlgorithmWatch. He is co-founder and publisher of the online magazine iRights.info (Grimme Online Award 2006). He has testified before several committees of the German Bundestag, i.e. on AI and robotics. Matthias serves on the governing board of the German section of Reporters Without Borders and the advisory councils of Stiftung Warentest and the Whistleblower Network and is a member of the steering committee of the German Internet Governance Forum (IGF-D). He has been a fellow of ZEIT Stiftung, Stiftung Mercator and the American Council on Germany. Matthias has written and edited books on digital journalism and Internet governance and was named one of 15 architects building the data-driven future by Silicon Republic in 2017. He holds master's degrees in Journalism from the University of Colorado in Boulder and in Philosophy from the Free University of Berlin.

Photo: Manuel Kinzer

/ **Marc Thümmler** / Publication coordinator

Marc Thümmler is in charge of public relations and fundraising at AlgorithmWatch. Prior to that he worked as a project manager for the German Museum for Film and Television Deutsche Kinemathek and the civil society organisation Gesicht Zeigen. Marc has extensive experience in film production as a production manager and editor, and he continues to develop and realise media in various formats as a freelancer.

ORGANISATIONS

/ AlgorithmWatch

AlgorithmWatch is a non-profit research and advocacy organisation, funded by private foundations and donations by individuals. Our mission is to evaluate and shed light on algorithmic decision-making processes that have a relevant impact on individuals and society, meaning they are used either to predict or prescribe human action or to make decisions automatically. We analyse the effects of algorithmic decision-making processes on human behaviour, point out ethical conflicts and explain the characteristics and effects of complex algorithmic decision-making processes to a general public. AlgorithmWatch serves as a platform linking experts from different cultures and disciplines focused on the study of algorithmic decision-making processes and their social impact; and in order to maximise the benefits of algorithmic decision-making processes for society, we assist in developing ideas and strategies to achieve intelligibility of these processes – with a mix of technologies, regulation, and suitable oversight institutions.
https://algorithmwatch.org/en/

/ Bertelsmann Stiftung

The Bertelsmann Stiftung works to promote social inclusion for everyone. It is committed to advancing this goal through programmes aimed at improving education, shaping democracy, advancing society, promoting health, vitalizing culture and strengthening economies. Through its activities, the Stiftung aims to encourage citizens to contribute to the common good. Founded in 1977 by Reinhard Mohn, the non-profit foundation holds the majority of shares in the Bertelsmann SE & Co. KGaA. The Bertelsmann Stiftung is a non-partisan, private operating foundation. With its "Ethics of Algorithms" project, the Bertelsmann Stiftung is taking a close look at the consequences of algorithmic decision-making in society with the goal of ensuring that these systems are used to serve society. The aim is to help inform and advance algorithmic systems that facilitate greater social inclusion. This involves committing to what is best for a society rather than what's technically possible – so that machine-informed decisions can best serve humankind.
https://www.bertelsmann-stiftung.de/en/