Chapter 19
The Issue of the Image of Algorithms

Lisa Käde

Algorithms have been a part of our daily lives for a long time. This paper will consider the need for a regulation of algorithms from three different perspectives of images.

Firstly, while they are usually associated with controlling machines, especially computers, algorithms are increasingly well suited to deal with images. In that capability, computers – enabled by algorithms – can help analyze large amounts of images at a time, generate entirely new images, sort them, and augment the human ability to perform tasks which require a great attention to detail. Secondly, images can also be used to label algorithms according to their functionality and impact. Visual representations of complex concepts can be an effective way to increase the transparency of algorithms and related processes. Thirdly, looking beyond the intuitive meaning of the word "image" as a visual representation, "image" also describes a mental impression or conception of something.[1] Appreciating how frequently humans are confronted with algorithms without knowing what exactly they are dealing with, it seems worth considering how the imagination influences the image of algorithms.

Still, when talking about algorithms and how to deal with them, dystopian – mostly science-fiction related – scenarios come to mind. We remember armies of supposedly friendly robots suddenly going awry, turning against humankind. We remember "PRECRIME" and its way to predict crimes before the culprits even conceive them.[2] And – above all – we think of artificial intelligence (AI) as an autonomous construct overcoming restrictions and surpassing human intelligence, killing everything that stands in its way.

However, people are not only concerned with robots in sci-fi movies. A very recent example of a seemingly discriminative algorithm, which

1 "Image", *Merriam-Webster.com Dictionary*, Merriam-Webster, https://www.merriam-webster.com/dictionary/image.

2 Referring to the movie "Minority Report", Twentieth Century Fox Film Corporation 2002.

selects thumbnail previews of pictures users post on *Twitter*, suggested unconscious racism on the platform. Thumbnail selections occur when an image is too large to be displayed in the respective context. Instead, a part of the picture is selected which seems to be representative of the picture. In this example, each image consisted of various pictures of different people. The algorithm seemed to prefer such parts of these pictures providing images of white people over such parts which depict people of color. Plenty of examples appeared to support this theory.[3] Still, it might not be just as bad as it seems, since such algorithms evaluate many factors in pictures to provide a satisfying result such as contrast, brightness, image quality etc. These are also factored into the Twitter algorithm's preview selection of images. If the images which display previously "discriminated" people are of better quality and contrast than the images of those people which the algorithm seemingly favoured before, suddenly the algorithm seems to neglect the latter while selecting the former as a thumbnail.

Another algorithm displaying discriminative behaviour made the news in 2015, when the Google Photo app introduced a feature which automatically labelled photos according to the content the algorithm recognized. A user posted a screenshot of how one of their friends was labeled as a gorilla.[4] While the problem probably lies in training data bias or imperfect automatic labelling of training data,[5] Google "resolved" this issue by simply removing the label "gorilla" altogether.[6] Moreover, Italy is considering using facial recognition and sound observation technologies in football stadiums to tackle issues of actual racism – leading to high-resolution images and sound recordings of conversations of visitors.[7] There was also intense discussion regarding the different COVID-19 contact tracing apps and data protection issues related to their use. These cases also highlight potential conflicts of interest and are therefore worth recalling when regulating algorithms.

---

3 Impressive examples are available at https://www.theguardian.com/technology/2020/sep/21/twitter-apologises-for-racist-image-cropping-algorithm. Meanwhile, Twitter reacted to this issue and promised to give users more control, see https://blog.twitter.com/en_us/topics/product/2020/transparency-image-cropping.

4 See, e.g., https://www.bbc.com/news/technology-33347866.

5 The algorithm might not have been presented with enough photos of people of colour, for example.

6 Rendering the program unable to detect actual gorillas, as well, see, e.g., https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/.

7 See https://algorithmenethik.de/2020/09/23/italien-in-echtzeit-gegen-rassistische-fussballfans.

The bad associations triggered by algorithms are some of the reasons why people may be cautious to implement algorithms more widely. One could say the (mental) *image* people have of algorithms seems to be an *issue*. Even though most of the truly terrifying scenarios are pure science-fiction (at least for now), they can still be overwhelming. "We cannot change what we do not understand"[8], says a character in a recent book by the author Schätzing, referring to the source code of a super intelligent system that evolved to a point at which humans could no longer control its actions. Most of the time we also avoid what we do not understand. A lack of understanding suggests a lack of control. And we don't like being out of control, it makes us feel helpless and fear sets in. Even though not all uses of algorithms might appear as drastic as the ones depicted in these scenarios, imposing regulation on supposedly dangerous subject matter seems to give us peace of mind.

It is therefore more than ever important to understand why – and how – those algorithms do what they do, to both improve them and calm down the discussing masses, and to prevent jumping to conclusions. To contribute to this understanding, this chapter will first analyze the need for algorithms to be regulated, taking into account the impact they have on society (I). After some remarks concerning terminology (of both algorithms as well as regulation) to limit the scope of this paper (II), existing approaches to regulation (of algorithms) are presented (III). Additionally, practical approaches to foster transparency and trust in algorithms will be briefly introduced (IV).

I. The Need for Regulating Algorithms

So, why do we need to regulate algorithms? Before answering this question, it should be noted that so far, it has not yet been specified what exactly is to be understood by an "algorithm". The reason is that the image of the abstract concept of "algorithms" is a very strong one and might give rise to a highly subjective understanding by different audiences. The meaning of the term will therefore be discussed only at a later stage. For now, this section will continue to embrace all different forms of algorithms.

The desire to regulate algorithms stems from various domains, including fear of algorithms or technology in general, biased data or algorithmic decision making, as well as the potential to improve software and interdis-

8 Schätzing (2018) 602 et seq.

ciplinary communication, and the awareness that algorithms are already impacting our everyday lives.

1. *Fear of algorithms and technology*

One might fear algorithms because of their opacity. The concept of the so-called "black box" – an algorithm which has input and output interfaces for interacting with a user, but does not offer insight into the inner workings, usually in the context of machine learning and AI – is often used to highlight algorithmic opacity. To many people, all kinds of algorithms are black boxes, simply because they cannot read code or understand how the program works. Artificial Intelligence is often referred to as a "black box"[9] which might fuel the apparent opacity of AI in general. Moreover, people might fear biased data, meaning machine learning algorithms trained on data selected by humans might inherently be subconsciously biased or discriminative.[10] In addition, some people are afraid of technology in general,[11] and fear that there is no way to ask a machine for clarification in the same way one could interact with a person.[12] The latter concerns especially decision-making systems when the decision made by the algorithm affects an individual's life.

2. *Improvements through regulation*

At the same time, regulation might present a chance to improve algorithms. Industry standards – which may result from regulation[13] – could guide developers to produce better code, and more thorough testing could potentially prevent damage once an algorithm is implemented and put to action. Regulation will also bring together people from many disciplines to come to acceptable terms for all parties.

---

9 De Streel et al. (2020) 3 et seq.; Pasquale (2015); Data Ethics Commission (2019) 189; German AI Strategy (2018) 16; European Commission (2018) 13.

10 See, e.g., Hajian/Bonchi/Castillo (2016); German AI Strategy (2018) 37; European Commission (2019) 18; Data Ethics Commission (2019) 167 et seq.

11 For research on "technophobia" see, e.g., Brosnan (2002) 10 et seq.

12 See results of representative phone interviews conducted by Kolany-Raiser/Heil/Orwat/Hoeren (2019) 15.

13 E.g. AlgoRules, https://algorules.org; industry standards already exist, e.g., for encryption, see Smid/Branstad (1988); Heron (2009).

3. Present impact of algorithms

Regulation is also not restricted to algorithms which might exist in the near or far future. Algorithms are already part of our daily lives. Users could encounter individual – possibly unfair – pricing in online shops or insurance rates,[14] there might be automated decision-making in the public administration for simple administrative acts,[15] recommendation systems already guide users through online shops and social media, possibly unconsciously affecting their behaviour. Others might be subject to an automated grant decision.[16] Many companies have algorithms pre-select their applicants, and some countries use software such as COMPAS to get recommendations for the early release of prisoners.[17]

As discussed at the beginning of this paper, the great power of AI in the context of image analysis also potentially poses risks of discrimination which should be addressed by regulation – both for prevention and mitigating effects on society.

All things considered, the topic of the regulation of algorithms seems like something that should have been dealt with a while ago. But, like most technology related aspects of regulation, the law is more reactive than anticipative of developments.

II. Some Remarks Concerning Terminology

Before one can dive into the discussion of regulating algorithms, some common ground should be found to clarify the basic terms. Even though there are many (abstract) ways to define both algorithms and regulation, no consensus seems to exist on a general definition on the term "algorithm",[18] while the definition of "regulation" merely seems to be depending on the context it is used in.

---

14 Paal (2019) 43 et seq.; Simon/Butscher (2001); Thomas (2012); see also the findings by the European Consumer Organisation (BEUC) in BEUC (2020).

15 Luthe (2017); Malgieri (2019).

16 On the issue of using algorithms to assess creditworthiness see, e.g., Data Ethics Commission (2019) 231.

17 See https://www.equivant.com/northpointe-risk-need-assessments; Brennan/Dieterich/Ehret (2009).

18 Künstner (2019) 36.

1. Regulation

According to *Wikipedia*, regulation is the "management of complex systems according to a set of rules and trends". Others define regulation as a means to "govern or direct according to rule" or "to bring order, method or uniformity".[19] In the legal context, regulation is usually used as a way to describe the process of imposing legal restrictions upon a subject matter. In EU law, "regulation" could be contrasted with the term "directive": while the former has binding legal force throughout EU member states, the latter needs to be implemented by national law.

2. Algorithm

The term "algorithm" could describe a "set of rules that precisely defines a sequence of operations"[20] or – in other words – an unambiguous instruction for the solution of a pre-defined problem. Notably, most definitions of algorithms steer clear of referencing specific *types* of algorithms, such as "machine learning algorithms" or "decision-making algorithms". The number of different algorithm species seems to be infinite.

Sometimes, in common language, computer programs in general are referred to as algorithms, whereas the term could also be used as an abstract description of a computer program and its underlying concepts (e.g., "the Google Search algorithm", "the Facebook newsfeed algorithm").[21] In the context of the regulation of algorithms, the term encompasses the abstract concept of automated processes as well as the specific issues of self-learning and self-improving systems.

The discussion on the regulation of algorithms seems to have received increasing attention over the past two to three years.[22] This is because research and applications of machine learning and AI are flourishing due to technical progress in hardware development and data availability. These kinds of algorithms are present in most of the above examples and are symptomatic for the "black box" discussion. Since the most recent publications of the EU and the German Federal Government on the regulation of

19 https://www.merriam-webster.com/dictionary/regulate.
20 Stone (1971) 4.
21 See also https://www.merriam-webster.com/dictionary/algorithm.
22 See GoogleTrends on "machine learning" and "artificial intelligence", https://trends.google.com/trends/explore?date=all&q=machine%20learning,artificial%20intelligence.

algorithms refer to AI and machine learning, this paper as well will focus on machine learning (ML) algorithms and AI.

3. *Regulating algorithms*

In a nutshell, one could describe the regulation of algorithms as *providing a legal framework for the development and use of algorithms.* But, as indicated before, the term "algorithm" is extremely broad. Even if it were restricted to computer programs, recurring to this definition would treat all algorithms equally, regardless of their complexity and purpose. While this might seem beneficial, the simple text editor "Notepad" does not need the same restrictions as complex systems like "COMPAS"[23] which eventually have legal effects on individuals. It should also be noted that developers could refrain from innovating and investors could stop providing funding simply because they fear contravening laws if regulation is too restrictive. Additionally, as fast as technology advances, there is no way to tell what kinds of algorithms we will encounter in the near future. A legal framework to regulate algorithms should therefore be flexible so that it only restricts those algorithms which need to be controlled in a way that is open to future developments. It should also, as precisely as possible, define which class of algorithms it strives to regulate, both because different algorithms pose different issues and to not inadvertently affect "innocent" algorithms.

Finally, it should be noted that regulation by way of a "legal framework" does not necessarily have to be comprised of formal laws. It could also include mandatory certifications, industry guidelines, EU directives and regulations. Self-regulation could also be factored in.[24]

III. *Examples of Existing Approaches to Regulation*

One way to respond to the issue of regulating algorithms is to consider the classification of algorithms: those which make up what is called "AI", for example, versus those implemented in domains such as online platforms like social media websites. The latter needs to regulate any kind of algo-

23 See footnote 16.

24 Künstner (2019) 40.; German AI Strategy (2018) 29; Data Ethics Commission (2019) 70 et seq. and 201 et seq.

rithm while considering the disparities of the platform and the users. The former deals with the specificities of a certain type of algorithm irrespective of its application, such as facial recognition, automatic thumbnail selection or deep fake image generation.

1. Initiating regulation through algorithm type-specific guidelines

In 2019, the EU High Level Expert Group (HLEG) on AI published its "Ethics Guidelines for Trustworthy AI".[25] Even though that does not sound like "regulating algorithms", a second – deeper – examination reveals some quite relevant thoughts which should at least guide the regulation of algorithms.

The aim of these guidelines is not to explicitly regulate, but to somehow encourage *trust* in algorithms – regulation might be one way of fostering such human trust. The HLEG identified four main ethical pillars which need to be addressed when dealing with AI (or algorithms in general):[26]

1) AI needs to always respect human autonomy
2) AI needs to always prevent harm
3) AI needs to be fair
4) AI needs to be explicable.

In a next step, the HLEG AI Guidelines drew up a non-exhaustive list of requirements for trustworthy AI which are in line with those ethical principles. This list includes human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental wellbeing, and accountability. Of course, these requirements and guidelines may sound convincing, but they are in no way binding for any programmer, company, or authority. They are also very abstract, and as such make no suggestions on how they could be incorporated into AI.

In a similar fashion, the German Commission on Data Ethics issued a report discussing data and algorithms, their impact on society and suggested ways of regulation.[27] The German Commission placed a great emphasis on human dignity, human autonomy as an expression of freedom, privacy, security (of privacy, goods, physical and emotional safety, environment),

25 European Commission (2018).
26 Ibid., 11 et seq.
27 Data Ethics Commission (2019).

democracy (digital technologies impact freedom of expression and freedom of information, among others), justice and solidarity, and sustainability (referring to the UN Sustainable Development Goals).[28]

The HLEG then elaborated on technical and non-technical methods to implement these requirements in practice, as did the German Commission on Data Ethics. This is an essential step towards an *actual* regulation that leaves the confinement of lengthy documents and is received by those individuals shaping and using the algorithms addressed by the regulation. Those suggestions include the establishment of certification mechanisms, standardisations, codes of conduct, thorough testing and validation, among others. However, one question remains: How can regulators reach the people designing and implementing AI and other algorithms?

2. *Regulating with respect to the domain of application*

Another approach of regulation does not address the supposed dangers of specific algorithms, but assesses the issues stemming from the situation where the algorithm is applied. An example would be the regulation of internet platforms, where users are dealing with an internet website interface, possibly providing personal or business data – either for delivery purposes, product display in online shops or even as payment, like in the case of ad-based services – without having any means of knowing what the algorithm will do with their data, or why they are shown the content they get to see. The issues of voter manipulation by means of tailored and manipulated (fake) news delivery come to mind.[29] TikTok users might wonder why they are presented with videos on specific topics, and some people of colour might question why automated towel or soap dispensers won't react to their activation gesture[30] or why virtual backgrounds in online conferencing tools do not recognize their faces.[31]

It is not as if the topic was a blank slate. There already are different kinds of statutes and EU directives and regulation pointing towards a regulation of algorithms, some of which will be discussed in this section.

---

28 Data Ethics Commission (2019) 43 et seq.

29 Referring to the *Cambridge Analytica* incident, see, e.g., https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election.

30 See, e.g., https://metro.co.uk/2017/07/13/racist-soap-dispensers-dont-work-for-black-people-6775909/.

31 https://twitter.com/colinmadland/status/1307111818981146626.

a) EU General Data Protection Regulation (GDPR)

Even though the GDPR does not contain the word "algorithm" or "computer program", it still deals with many topics and situations related to algorithms. A search for the word "automated" will lead to several recitals and provisions which are dealing with the issue of regulating algorithms – including, but not limited to, recitals 15, 67, 68 and 79 as well as articles 2, 20, 21 and 22.

Prima facie, data protection seems to be a topic of static information which might be stored digitally. But – obviously – data protection nowadays is above all concerned with *algorithms* that store data, considering that data subjects do not have full access to the algorithms involved, transparency is usually difficult to provide, and algorithms make it easy to deal with great amounts of data at very low cost. Thus, at the very beginning of the regulation, in Article 2.1, the GDPR limits its scope to the "processing of personal data wholly or partly by automated means and to the processing other than by automated means […] which form part of a filing system or are intended to form part of a filing system."

Article 22 deals with the issue of automated individual decision-making, even though the topic is not necessarily data protection related. It reads: "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" (emphasis added). This section provides succinct criteria which – if fulfilled – not only afford the data subject with a right not to be subject to such a decision (except for the cases stated in paragraph 2, which include situations in which the decision is necessary for entering into or performing a contract between the data subject and the data controller), but also poses a duty to "implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests" in paragraph 3. Of course, there are many ways to circumvent this provision, by having humans confirm the automated decision, for example. Moreover, it indicates what the EU deems worthy of regulating: The focus is not on a certain algorithm, but specific real-life situations.

When the GDPR was first introduced, due to the high fines imposed by it, many companies (and individuals, as well) invested time and money to ensure that their data processing was transparent (e.g., by providing data privacy statements on their websites). As a regulation with direct impact on the EU member countries, the GDPR at least achieved some degree of transparency of data processing. Technically, it does not regulate *algorithms* but those who apply algorithms in their processing of personal

data. It does therefore not try to influence the structure or development of algorithms, but manage the effect on those affected by the algorithms – it could be regarded as a regulation of the *use of algorithms*.

b) Ranking: regulation in a B2B context

Algorithms are also widely used to rank goods and services on online platforms and search engines. Since these platforms are the basis of many marketing and sales concepts and algorithms used in that context have the potential to influence competition to a great degree, regulation seems appropriate and necessary to provide transparency. One approach regarding the issue of ranking goods or services online in a B2B context has recently been published in the "Guidelines on ranking transparency" by the EU, pursuant *to* Regulation (EU) 2019/1150[32], aiming to protect not consumers but providers of goods and services which rely on online intermediate providers to present their products for sale.

The regulation as well as the guidelines take into account the power and visibility of high-ranked goods and services in search engines and other online platforms, such as online warehouses, and require those intermediate providers to transparently explain their ranking mechanisms. These include the parameters used to rank entries, and the guidelines explicitly state that they apply irrespective of the technologies used for ranking.[33] The guidelines were published to support providers of online intermediation services and search engines in being compliant with Regulation 2019/1150. The regulation itself is legally binding in the EU member states, therefore – contrary to the guidelines on AI – these guidelines are less abstract and (probably) more relevant to those implementing and using algorithms for their purposes.

32 Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services, O.J. EU L 186, 57 et seq.

33 "Individual assessment and technologically neutral approach", section 1.3.2 of the Guidelines.

c) Automated administrative acts in Germany

In an entirely different area of law, Germany has a provision on "entirely automated acts of administration" (in the context of social administration). § 31a Social Code (SGB) Book X reads: "An administrative act can be entirely produced by automated facilities, as long as there is no need to have the individual case processed by an official. If the authority uses automated facilities to produce administrative acts, it has to appreciate such actual information relevant to the individual case provided by the person concerned which would not be determined in the automated process".[34]

Like the GDPR, the section does not regulate a specific algorithm, but rather uses the more general term "automated facilities" and thereby regulates any administrative act which is not produced by a human being. Notably, the provision does not require any kind of outward transparency. § 31a SGB X aims at providing technology-neutral electronic administrative services.[35] It does not alter the existing provisions on administrative acts but aims at ensuring that subjects of administrative acts are not disadvantaged by the automation of said acts.[36]

A similar provision exists for administrative acts in general, see § 35a Administrative Procedures Act (VwVfG). It restricts the use of automation to situations where there is no room for evaluation regarding the decision of the respective administrative act. § 35a VwVfG also requires that such automated decision be allowed by an applicable law, to ensure that only suitable procedures are making use of automation.[37] This could include the images taken by automated speed cameras which result in speeding tickets being automatically sent to the respective individuals. This is not yet being practiced in Germany.

3. Summary

These were only a few examples of laws which are already regulating (the use of) algorithms. More can be found in the regulation of algorithmic trading in the German Securities Trading Act (§ 80 II ff.) or article 18 of

34 Translation by the author.
35 Heße, Sabine, 'Commentary on § 31a SGB X', in: BeckOK Sozialrecht, note 2.
36 Ibid. note 5.
37 Luthe (2017).

the EU Directive on markets in financial instruments, for example,[38] and even more are in the making.[39]

There are many different legal areas trying to address the issue of automation and algorithmic involvement. Each of them deals with different issues and treats them in different ways. For developers and those responsible for IT-systems and applications, this would mean that they must be aware of all laws possibly applicable to an algorithm they are developing to be compliant. This becomes more complicated if multi-purpose algorithms are involved, meaning that upon creation, it is unclear how those algorithms will be deployed (such as many machine learning algorithms).

Discussions in Germany frequently address the idea of creating an "Algorithmen-TÜV", which means creating an institution responsible for testing algorithms.[40] However, this idea doesn't seem promising for several reasons. Firstly, it would be a German solution to an international problem. After all, algorithms are created everywhere and in uncountable numbers. There is no way a single national institution could thoroughly check all of them in a timely manner. While one could claim that such an institution could either only check algorithms used by the national administration or provide a general list of approved algorithms, this approach does not seem to be able to keep pace with the speed of algorithm development and could hinder innovation and digitalisation. Secondly, a certification like a TÜV-seal could suggest false confidence in new technology to the likes of the Diesel Scandal. How would one define which algorithms need to be certified? Also, the certification would most likely require companies to provide their source codes for the certifiers to "look into the black box". This again could hinder innovation if companies cannot provide their source codes due to contractual obligations or trade secret considerations and would thus rather avoid implementing algorithms which are subject to certification.

---

38 Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU, O.J. EU L 173, 349 et seq.

39 E.g., the Digital Services Act, see https://www.euractiv.com/wp-content/uploads/sites/2/2020/12/Digital_Services_Act__1__watermark-3.pdf and the Digital Markets Act, see https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2349.

40 See, e.g., https://algorithmenethik.de/2017/09/11/was-die-wahlprogramme-ueber-maschinen-sagen-die-menschen-bewerten/ and https://www.spd.de/fileadmin/Dokumente/Regierungsprogramm/SPD_Regierungsprogramm_BTW_2017_A5_RZ_WEB.pdf, 73.

To date, there is no known concept for the implementation of this kind of certification. Instead, the German Federal Government created an AI Observatory which is supposed to analyse potential effects and risks of AI in the context of work and society[41] which leans towards the solution discussed in this paper: Regulating specific situations or application contexts,[42] not specific types of algorithms.[43]

Algorithms of AI (mostly machine learning algorithms) for example, tend to be complex structures and their outcomes can be hard to predict, especially for non-machine-learning-specialists. Moreover, there is not just one kind of AI algorithm. Developers and data scientists frequently are coming up with new approaches and applications for machine learning. Regulating with regard to situations and environments could therefore be especially helpful when dealing with AI, since it does not need the regulator to anticipate future developments, but instead shifts responsibilities to the makers and creators of such algorithms. These are then required to ensure that their systems meet transparency and accountability requirements.

The regulations described above show how this aim can be achieved. It is not advised to impose a general "law on AI", since this would, on the one hand, be confined to technology as-is, and on the other hand, might be circumvented by the use of algorithms which are similarly damaging, but which do not fall under the definition of AI, if there even is one. The European Commission White Paper on Artificial Intelligence suggests creating a regulatory framework which should be applicable to all products and services making use of AI,[44] and then dives into the issue of defining AI. This could be avoided by taking the situation-centered or application-centered approach, therefore taking into account the effect algorithms have, irrespective of their type or implementation. The EC aims at developing a risk-based approach,[45] but restricts this risk-based approach to AI. It is questionable whether this pre-selection of algorithms is necessary.

Nevertheless, situation specific regulation also has its limits in that it seems impossible to address all situations individually (e.g., face recognition can have various applications: supporting immigration agents, identifying fugitives in large crowds, detecting people in traffic situations or

41 See https://www.denkfabrik-bmas.de/en/projects/ai-observatory.
42 See also AI Ethics Impact Group (2020) 35.
43 For the suggestion of a situation specific requirements' matrix as a basis for regulation, see Martini (2019) table at 76.
44 European Commission (2020) 16.
45 Ibid. 17.

tagging friends in picture collections, just to name a few). Such regulation therefore needs to carefully consider broader contexts in which such situations potentially arise, such as the GDPR is dealing with all situations where personal data is being automatically processed.

IV. Improving the Image of Algorithms Outside of Legal Regulation

Laws are not the only means of guiding the development and use of algorithms. The image – as in the public perception – of an algorithm might already be improved by voluntarily providing information on the use and functionality of algorithms. This section will present two suggestions of such rather self-regulative measures.

1. Algo.Rules

The first of these non-regulatory schemes is "Algo.Rules"[46] by the German group algorithmenethik.de in cooperation with irights.lab. These rules were created in a joint conversation that involved almost 500 participants from the areas of science and research, industries and organisations, civilians, NGOs, politics and administration. "Algo.Rules" provide guidelines on how to incorporate these rules into algorithmic projects, including detailed questionnaires.

To just point out of some of them: The very first rule would be to strengthen competency, addressing the issue that decision makers and developers alike need to understand both the functioning of the algorithm as well as the effects it could have when put into practice. In addition, safeguarding manageability addresses the issue of the algorithm staying in control of a human, something which we already saw in the HLEGs requirements of trustworthy AI. Moreover, ensuring intelligibility – like the requirement of explicability – tries to manoeuvre the algorithm away from being regarded as a black box into the direction of understandable and explainable decisions.

With regard to the above suggested situation-specific regulation, these rules are a tool to guide developers in creating algorithms which comply with said regulation. The rules heavily focus on anticipating effects, en-

46 All of those rules are described in-depth and accompanied by practical recommendations online at https://algorules.org.

suring transparency and maintaining accountability and therefore complement situation-specific regulation.

2. *Google model cards*

The second suggestion is a visualising approach, i.e., using images to improve the image of their algorithms. This is similar to food labels informing consumers about the contents and nutrition facts. In a similar way, algorithms could be labelled to give users a quick overview of how the algorithm functions, what its limitations are, and perhaps even provide the means to easily test it on their own data. Of course, simply stating that whatever system has "AI inside" or even naming the machine learning algorithm used, is not guaranteed to provide transparency to users, since they might not be aware of the features of specific algorithms, or the effect of automatic processing in general.

Google suggested these so-called "model cards"[47] for systems using machine learning models.[48] Two model cards are currently available, one for a model on Face Detection and one for Object Detection, both of which target machine learning algorithms detecting face and objects in images The model cards describe what kind of input a model requires, the expected outcome (e.g., whether the model will highlight the area of the input image which lead the model to detect a face), and its limitations. It is more a proof of concept now than an established mechanism, but it is a concise suggestion which could be integrated into a software development process if it reaches a status of "best practice".[49]

47 See https://modelcards.withgoogle.com/about.

48 In the context of machine learning, models are the trained or trainable structures used for making predictions or generating creative output. These models are used in the context of the executing computer program and roughly represent the underlying statistical algorithm.

49 A similar approach was suggested by GI (2020), https://gi-radar.de/276-beipackzettel-fuer-ki/; also demanding an obligatory AI label https://www.n-tv.de/panorama/Maschinen-ueberwinden-Schreibblockade-article22201094.html?utm_source=pocket-newtab-global-de-DE.

V. What Comes Next?

So, what comes next? Do we need a centralized "code for code",[50] unifying all of those far-spread different sprinkles of regulations of algorithms? For example, the Data Ethics commission suggested a "horizontal basic rule by means of an EU directive for algorithmic systems" on the European level to be accompanied by sector specific national legislation.[51] According to the German AI Enquete Commission, sector specific legislation might then be extended by AI specific provisions.[52] This seems like a reasonable approach: Identifying relevant sectors, situations or applications of algorithms, and then – if necessary – enriching regulation by provisions taking into account potential specific issues of AI. In this way, practitioners especially in the software development business could focus on legislation pertaining to their domain of application, without the need to assemble fragments of regulations according to their use of algorithms.

It should also be discussed whether industry guidelines, putting money into certifications (who do we trust to issue such certifications?), or laws are the preferred means of regulation. How much control is wanted and needed? How much responsibility is desired and required – and who should be responsible at all? How can one steer clear of over-regulation, taking into consideration constitutional rights such as freedom of opinion?

In conclusion, the image of algorithms might be improved by strict regulation, insofar as subjects to algorithms increasingly trust the legislators in protecting them from potential harm. However, while this supposed trust might seem to be comforting, it does not change the image of algorithms per se, since demonstrating a strict regulative approach might even emphasize the dangers and threats associated with the use of algorithms.

One should be aware of the potential manipulative and sometimes unconscious effects algorithms might have both on an individual's life and on democracy. But computer literacy might also go a long way, enabling users to better understand what potential threats they might be faced with, thus raising awareness and addressing the issue bottom-up in addition to the top-down approach of regulation.

---

50 Discussing a "Lex Algorithmica" in France, see GI (2018) 113.

51 Data Ethics Commission (2019) 180

52 See the summary of the German AI Enquete Commission report at https://dserver.bundestag.de/btd/19/237/1923700.pdf.

References

AI Ethics Impact Group (2020): 'From Principles to Practice – An interdisciplinary framework to operationalise AI ethics' (2020; https://www.ai-ethics-impact.org/resource/blob/1961130/c6db9894ee73aefa489d6249f5ee2b9f/aieig---report---download-hb-data.pdf)

BEUC (2020): 'The Use of Big Data and Artificial Intelligence in Insurance' (2020; https://www.beuc.eu/publications/beuc-x-2020-039_beuc_position_paper_big_data_and_ai_in_insurances.pdf)

Brennan, Tim/Dieterich, William/Ehret, Beate (2009): 'Evaluating the predictive validity of the COMPAS risk and needs assessment system', 36 Criminal Justice and Behavior, No. 1 (January 2009) 21–40

Brosnan, Mark J. (2002): *Technophobia – The Psychological Impact of Information Technology* (London: Routledge 2002)

Data Ethics Commission of the Federal Government of Germany (2019): 'Opinion of the Data Ethics Commission' (2019; https://www.bmjv.de/SharedDocs/Downloads/DE/Themen/Fokusthemen/Gutachten_DEK_EN.pdf?__blob=publicationFile&v=1)

De Streel, Alexandre/Bibal, Adrien/Frenay, Benoit/Lognoul, Michael (2020): 'Explaining the Black Box: When Law Controls AI' (2020; https://cerre.eu/wp-content/uploads/2020/03/issue_paper_explaining_the_black_box_when_law_controls_ai.pdf)

European Commission (2018): High Level Expert Group on Artificial Intelligence: *Ethics Guidelines For Trustworthy AI* (2018, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419)

European Commission (2020): 'Whitepaper on Artificial Intelligence – A European approach to excellence and trust' (2020; https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf)

Federal Government of Germany (2018): 'Artificial Intelligence Strategy' (2018; https://www.ki-strategie-deutschland.de/home.html?file=files/downloads/Nationale_KI-Strategie_engl.pdf), cited as: German AI Strategy

Gesellschaft für Informatik (2018): 'Technische und rechtliche Betrachtungen algorithmischer Entscheidungsverfahren – Studien und Gutachten im Auftrag des Sachverständigenrats für Verbraucherfragen' (2018; https://www.svr-verbraucherfragen.de/wp-content/uploads/GI_Studie_Algorithmenregulierung.pdf)

Hajian, Sara/Bonchi, Francesco/Castillo, Carlosn(2016): 'Algorithmic Bias – From Discrimination Discovery to Fairness-Aware Data Mining', KDD '16 – Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2125–2126 (2016; https://dl.acm.org/doi/pdf/10.1145/2939672.2945386)

Heron, Simon (2009): 'Advanced Encryption Standard (AES)', Network Security (2009) 8–12

Kolany-Raiser, Barbara/Heil, Reinhard/Orwat, Carsten/Hoeren, Thomas (2019): *Big Data* (München: C.H. Beck 2019)

Künstner, Kim Manuel (2019): 'Preissetzung durch Algorithmen als Herausforderung des Kartellrechts', Gewerblicher Rechtsschutz und Urheberrecht (2019) 36–42

Luthe, Ernst-Wilhelm (2017): 'Der vollständig automatisierte Erlass eines Verwaltungsaktes nach § 31a SGB X', Die Sozialgerichtsbarkeit (2017) 250–258

Malgieri, Gianclaudio (2019): 'Automated Decision-Making in the EU Member States – The right to explanation and other "suitable safeguards" in the national legislations', 35 Computer Law & Security Review (2019) 2–26 (https://reader.elsevier.com/reader/sd/pii/S0267364918303753?token=5D8EBF5DD9AFA9741959E608A623E0693A93E5DDFF29D24384EFB9A93DB4670DF78A331FEF74DCE05DEB4EBE8503613)

Martini, Mario (2019): 'Fundamentals of a Regulatory System for Algorithm-based Processes' (2019; https://www.vzbv.de/sites/default/files/downloads/2019/07/19/martini_regulatory_system_algorithm_based_processes.pdf)

Paal, Boris (2019): 'Missbrauchstatbestand und Algorithmic Pricing', Gewerblicher Rechtsschutz und Urheberrecht (2019) 43–53

Pasquale, Frank (2015): *The Black Box Society. The Secret Algorithms that Control Money and Information* (Cambridge, Mass.: Harvard University Press 2015)

Schätzing, Frank (2018): *Die Tyrannei des Schmetterlings* (Cologne: Kiepenheuer & Witsch 2018)

Simon, Herrmann/Butscher, Stephan A. (2001): 'Individualised Pricing: Boosting Profitability with the Higher Art of Power Pricing', 19 European Management Journal, Issue 2 (2001) 109–114

Smid, M.E./Branstad, D.K. (1988): 'Data Encryption Standard – Past and Future', 76 Proceedings of the IEEE, Issue 5 (1988) 550–559

Stone, Harold S. (1971): *Introduction to Computer Organization and Data Structures* (New York et al.: McGraw-Hill 1971)

Thomas, R. Guy (2012): 'Non-Risk Price Discrimination in Insurance – Market Outcomes and Public Policy', 37 The Geneva Papers on Risk and Insurance – Issues and Practice (2012) 27–46