WHITE PAPER ON ARTIFICIAL INTELLIGENCE

Proposals of the Federation of German Consumer Organisations - vzbv

11. May 2020

Impressum

Verbraucherzentrale
Bundesverband e.V.

Team
Digitales und Medien

Rudi-Dutschke-Straße 17
10969 Berlin

digitales@vzbv.de

Bundesverband der Verbraucherzentralen und Verbraucherverbände

Verbraucherzentrale Bundesverband e.V.

CONTENT

I. SUMMARY

Overall Verbraucherzentrale Bundesverband (vzbv) welcomes the proposal of the European Commission for a future regulation of Artificial Intelligence (AI). The White Paper proposes meaningful approaches. However, vzbv holds that the White Paper is in some respects not ambitious enough and the scope of the proposed regulation is too narrow, not covering a number of high-risk AI applications.

The definition of the risk and potential damage of AI is too narrow

The definition of the risk and potential damage of AI applications rest only on effects they have on individuals/legal entities. This is too narrow. The definition of high-risk AI applications and the harm should include damages to the society at large and to groups. Monetary & economic losses as well as losses due to denied access to markets/scarce resources should be mentioned explicitly not only implicitly as “material damage.

Many high-risk applications fall out of the scope of legal requirements

According to the White Paper, legal requirements shall only be imposed for high-risk applications if they fulfil both criteria (cumulative approach):

- The application is in a predefined high-risk sector (defined by an exhaustive list: Healthcare, transport, police, judiciary)
- The application itself poses a high risk

However, it is vzbv’s position that applications in any sector must be subject to legal requirements if they pose a high risk/can cause significant damage (e.g. insurance, discrimination in e-commerce or smart digital assistants).

The risk-assessment system proposed by the White Paper is binary: Applications are either high-risk (legal requirements apply) or low-risk (no legal requirements). Risk and regulatory requirements should be assigned gradually in different levels. Legal requirements should reflect the corresponding risk level.

Too little transparency for consumers

Mandatory rules for transparency towards Consumers: So far, the White Paper only foresees information in form of labelling of high-risk AI so that users know that they interact with an AI. This is too narrow. Consumers need to know about the risks, data base etc. Developers and operators of AI must be able to explain how their systems work to ensure traceability (and accountability). Data subjects must be provided with all the information necessary to exercise their rights when necessary.

II. INTRODUCTION

The shift towards algorithm-based decision making (ADM) and artificial intelligence (AI) is changing the way in which consumer markets and our societies function. vzbv welcomes that the European Commission acknowledges that AI will have a major impact on the welfare of individuals and society at large and that it will play an increased role in the entire socio-economic sphere.

Therefore, vzbv supports the European Commission’s effort to introduce a risk-based European regulatory framework for AI in order to reap its benefits and minimise the risks associated with it. vzbv stresses that the European Union must take responsibility as a global actor and establish a regulatory framework on AI that ensures people can trust that the technology respects European values and European laws.

EU’s responsibility for driving global AI rules

Due to a lack of ambition for a human-centred regulatory approach from the USA (due to a political stalemate in Washington) or from China, the European Union currently is the only major political block that is capable of driving forward worldwide rules for a technology in order to place the human being and human values in the centre of its vision. vzbv is convinced that a European regulatory framework on AI has the potential to act as a global standard and beacon, and model for other administrations around the globe, similar to GDPR.

The EU must accept its responsibility and historic role as a regulator for AI whose decisions will shape the trajectory of technological development worldwide and in the future. In light of this historic role, vzbv stresses that the White Paper on AI introduced the right approaches for an effective regulatory framework. Nonetheless, it needs amendments as the scope of its application is too narrow and the proposed legal requirements are insufficient.

Consumer trust as a driver for AI adoption

Currently, many consumers – rightly – distrust AI, largely due to the obscure nature through which decisions are made or prepared. This level of distrust hampers the adoption and demand for AI in the European Union (EU).

A European regulatory framework on AI must ensure that consumers can trust that AI adheres to European laws and values in order to encourage uptake and dissemination of the technology.

The framework must ensure that systems are controlled effectively by independent auditors, AI systems making/preparing decisions on consumers adhere to high quality standards, are highly transparent for consumers and enables consumers to exercise their legal rights. This will act as a driver for the development of a European AI industry: Trust in the systems encourages consumers to adopt AI, thereby driving demand for AI made in the EU and dissemination of the technology through all social and economic spheres.

vzbv's comments will refer to the White Paper "On Artificial Intelligence - A European approach to excellence and trust" (COM (2020) 65 final from 19 February 2020), referred to as "The White Paper" in the following.

III. COMMENTS ON THE WHITE PAPER ON AI AND PROPOSALS FOR A REGULATORY FRAMEWORK

1. A REGULATORY FRAMEWORK FOR AI MUST COMPLEMENT GDPR: PROFILING, SCORING AND PREPARATION OF HUMAN DECISIONS

The European Commission notes that the feedback on the guidelines of the High-Level Expert Group on AI revealed that requirements "regarding transparency, traceability and human oversight are not specifically covered under current legislation in many economic sectors." In contrast, the White Paper refers to the GDPR on several occasions in order to point out that consumer rights in the realm of personal data – e.g. with respect to transparency of AI decisions – are already covered by existing legislation and no further legislation is needed.

vzbv, in line with the German Data Ethics Commission (DEK) and the academic community[1], points out that it is not sufficient to simply refer to GDPR when it comes to the

[1] *Data Ethics Commission*, '*Opinion of the Data Ethics Commission*', 2019 <https://www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_EN_node.html>. *Mario Martini*, *Fundamentals of a Regulatory System for Algorithm-Based Processes - Expert Opinion Prepared on Behalf of the Federation of German Consumer Organisations (Verbraucherzentrale Bundesverband)*, 2019 <https://www.vzbv.de/sites/default/files/downloads/2019/07/19/martini_regulatory_system_algorithm_based_processes.pdf>.

protection of personal data in the context of AI: the GDPR's scope is limited and does for example not regulate profiling or scoring per se or the automated preparation of human decisions. One core function of AI applications in consumer markets is the classification and prediction of user behaviour based on profiles/scores in order to prepare or make/prepare decisions about consumers. Furthermore, GDPR only covers personal data. However, AI applications increasingly rely on non-personal data, also when preparing or making decisions about consumers, which leaves consumers unprotected. It also becomes increasingly difficult to clearly distinguish personal and non-personal data. This underpins the urgent need to supplement GDPR with specific rules on profiling/scoring and automated preparation of human decisions.

> vzbv therefore suggests that the European Commission complements the GDPR with a regulatory framework on AI:
>
> It must define minimum legal requirements for profiling and scoring: Profiling as such (and not just decisions based on it, as defined in GDPR) must be regulated. In this respect, red lines should be defined, requirements should be standardised in order to clearly determine lawfulness of applications, and the principle of proportionality should be specified.
>
> The framework must also establish legal requirements related to the quality of (partially) automated decision-making and profiling: In order to ensure that the results of decisions are legal and correct, substantive procedural requirements should be laid down. For example, the data used for the decision should only be processed using recognised mathematical-statistical methods and the data should be demonstrably relevant to the decision in question.

2. SCOPE OF THE REGULATORY FRAMEWORK

vzbv welcomes the European Commission's idea of introducing a risk-based approach to regulate AI. However, vzbv thinks that the scope of the approach described in the White Paper is too narrow and needs several amendments.

2.1 The definition of AI

vzbv realises that there are many definitions of what AI is, so that the term "AI" represents a concept that is inherently ambiguous and not objectively defined.

The definition of AI proposed in the related section of the White Paper is unclear. In Footnote 47 (p. 16) the European Commission refers to the definition of AI of the High-Level Expert Group on AI (HLEG). However, the HLEG's definition of AI seems to be too narrow and focused on machine learning models as it states that "AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions." This definition runs danger of being too narrow, as not all systems relevant to this regulatory framework are able to "adapt their behaviour by analysing how the environment is affected by their previous actions." As the regulatory framework should be technology-neutral, the reference to the machine learning model in the definition of AI unnecessarily narrows down the scope of this framework. It excludes systems that do not adapt their behaviour depending on how their previous actions affected the environment. Also, it is possible that systems become excluded from the scope whose adaptation happens at greater intervals.

On the other hand, in an attempt to provide a simple and straightforward definition, the European Commission states that: “Simply put, AI is a collection of technologies that combine data, algorithms and computing power” (p. 2). This definition could apply to any software, which would be beyond the scope of the White Paper.

The European Commission seems to be oscillating between a very detailed but too narrow definition of AI (HLEG) and a very simple, but too broad definition of AI.

Choosing a definition of AI is crucial and has major consequences for the regulatory framework. If the European Commission adopts such a narrow definition, the application of legal requirements could be questioned by simply arguing that a specific application cannot be considered as AI. Such a narrow definition would allow companies/organizations to easily bypass and circumvent future rules. This would ultimately lead to a lack of accountability which undermines citizens trust in AI.

The European Commission should use a term/definition of AI that is valid, comprehensive and encompasses present and future applications that are potentially posing risks for individuals (persons or legal entities), for society at large and for specific social groups.

‘Algorithm-based decision-making’ (ADM hereinafter) on the other hand is a technology-neutral concept. It includes the technologies that the HLEG and the public generally refer to as “Artificial Intelligence”. This is the reason why the German Data Ethics Commission also worked with the term “Algorithmic Systems”[2]. Even if the European Commission chooses to stick with the term “AI” it should use the definition of “algorithmic-based decision-making” to define the technology that it targets with its regulatory framework:

ADM comprises much more than just code or an algorithm. It refers to the entire process from data acquisition and data analysis to the interpretation of the results and the way to derive a decision or a decision recommendation from the results[3]. ADM are characterised by the fact that they contain an algorithmic component (“control system”) which produces an output (“decision”) on the basis of an input, and outputs it in the form of a (numerical) value. As such, ADM also includes “learning” systems that derive decision-making rules from data by means of machine-learning and can adapt them over time. The systems discussed under the keyword Artificial Intelligence (AI) usually fall under this definition. The key element of the term “ADM” is relevant for policymaking as it stresses the element that the system produces an output that is used to prepare or take a decision that has an impact on people or legal entities.

> vzbv proposes to use a more specific term, not reducing everything to ‘AI’. It seems that the use of the concept of ‘algorithmic-based decision-making’ (ADM) is more aligned with the regulatory desires of the White Paper and more suitable to define its action field.

[2] Data Ethics Commission.p. 59-62, p. 160.

[3] See: *Kilian Vieth and Ben Wagner*, *Teilhabe, Ausgerechnet - Wie Algorithmische Prozesse Teilhabechancen Beeinflussen Können*, 2017 <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/teilhabe-ausgerechnet>.

2.2 The definition of risk and damage: Consider harm for society and social groups

The current assessment of the level of risk of AI applications seems to consider only possible damages to individual entities (individual persons or legal entities[4]).

This is in stark contrast to the introductory remarks of the White Paper that rightly states that "*the impact of AI systems should be considered not only from an individual perspective, but also from the perspective of society as a whole*"[5].

It is not obvious why effects on social groups should not be considered. If AI applications have a negative impact on many individuals of a specific group, the individual harm may be small, but the accumulated loss of welfare could be very high (e.g. through price differentiation or denial of certain levels of service).

AI can lead to people being treated differently according to group characteristics that are not (directly or indirectly) protected by antidiscrimination law, thereby reinforcing existing social inequalities. The fact that specific groups are not protected by existing anti-discrimination law highlights the importance to extend the scope of application of the AI regulation to social groups.

Example: Price discrimination and group targeting

For example, price discrimination can cause loss of welfare for specific social groups like single parents. The monetary effect might be small for the individual but large at the aggregated group level and it can undermine trust in the market economy. AI can be employed to steer people's access to certain markets or market segments, e.g. the job or housing markets. AI applications for targeted advertisement could (deliberately or not) exclude people of lower income from accessing higher segments in the job or housing market, thereby reinforcing existing social divides by preventing upward social mobility[6]. Also so-called fake news can undermine trust in democracy, thereby undermining the social fabric in general.

> vzbv recommends that the European Commission explicitly states in the definition of the scope of the future regulatory framework that the assessment of the risk level of AI applications must include the impact on members of specific social groups and society at large.

2.3 A horizontal definition of high-risk applications is needed - no confinement to specific sectors

The White Paper proposes to introduce legal requirements only for high-risk applications. According to the White Paper however, applications are only high-risk, if they fulfil the two criteria below (cumulative approach):

[4] *Europäische Kommission*, *Weißbuch Zur Künstlichen Intelligenz – Ein Europäisches Konzept Für Exzellenz Und Vertrauen*, *COM(2020) 65 final*, 2020 <https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_de.pdf>., p. 17: *"The assessment of the level of risk of a given use could be based on the impact on the affected parties. For instance, uses of AI applications that produce legal or similarly significant effects for the rights of an individual or a company."*

[5] Europäische Kommission. p.2

[6] For example of how advertisement can be discriminatory compare: *U.S. Department of Housing and Urban Development (HUD)*, '*HUD charges Facebook with housing discrimination over company's targeted advertising practices*', 2019 <https://www.hud.gov/press/press_releases_media_advisories/HUD_No_19_035>.

- Applications are in a predefined high-risk sector (defined by an exhaustive list including healthcare, transport, energy and parts of the public sector).
- The application itself poses a high-risk.

vzbv holds that this approach is to narrow in scope. The White Paper lists healthcare, transport, energy and parts of the public sector as possible high-risk sectors. This is understandable as these sectors inherently present high risks for different reasons.

But there are also **many high-risk applications outside** of these **predefined sectors** that can have strong negative impact of different nature on society, groups and individuals. If the approach of the White Paper is maintained, these would fall out of the scope of the AI regulation (see examples below).

Example: Personality analysis and consumer background checks to control access to markets/services is high-risk

Analysis of personality traits based on online data got infamous in the wake of the Cambridge Analytica scandal, where it was used for political campaign management[7]. Online platforms like Airbnb can employ personality analysis as a way to classify users in order to determine which user gets access to which service/platform/market and under what conditions[8]. The downside is the very possible, unjustified exclusion of (groups of) users from entire markets or market segments.

The economic damage due to misclassification of users might amount to a large loss of welfare– even in individual cases when users are denied access to the platform due to biased training data, erroneous algorithmic inferences or data records. The aggregated welfare loss on a collective level might amount to millions of euros for specific groups that are systematically excluded or put at a disadvantage.

A critical example that would fall outside of the scope of the proposed regulation if the risk definition of the White Paper is maintained includes different kinds of AI-based personality trait analysis employed in order to control access to markets or services (this holds as well for other AI-based background checks on consumers and consumer reporting services[9]): Personality analysis might happen in real time based on biometric data or on data found on the web (e.g. social media data). For example, Airbnb holds a patent aiming to conduct personality analysis based on online/social media data in order to determine the trustworthiness and risk of its users and to optimise the matching process between hosts and guests. Criteria to classify users according to trustworthiness or risk include among others alleged drinking habits, alleged involvement in pornography or authoring of “online content with negative language”, probably deducted from social media. It also includes “social connections”, employment and education record[10]. It is not known in how far Airbnb employs this software yet. Still, it is obvious that classifying users based on such data is bound to be faulty on a regular basis: Restaurant owners advertising

[7] *The Guardian*, ‘*Cambridge Analytica: how did it turn clicks into votes?*’, May 2018 <https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie>.

[8] *Evening Standard*, ‘*Booker beware: Airbnb can scan your online life to see if you’re a suitable guest*’, 2020 <https://www.standard.co.uk/tech/airbnb-software-scan-online-life-suitable-guest-a4325551.html>.

[9] *Vox*, ‘*Beware of these futuristic background checks*’, 2020 <https://www.vox.com/recode/2020/5/11/21166291/artificial-intelligence-ai-background-check-checkr-fama>.

[10] Evening Standard.

vineries whose wine they serve might be wrongly classified as alcoholics. Journalists investigating illegal trafficking of forced sexworkers might be wrongly classified as being involved in pornography. People with lower education might find it difficult to find apartments in areas where hosts have a higher educational background (e.g. in inner cities).

Example: AI-based reimbursement check in health and car insurance

Some private health insurances manage the approval of reimbursement claims digitally: Customers enter the received treatment in the online form, upload the invoice and click on “Pay invoice”. The assessment whether the treatment is covered by the terms of the tariff of the specific contract is performed digitally.

In the car insurance sector, AI is employed to examine photos of damages and to decide on the coverage of the damage or repair costs. For instance, fully automated processing (damage assessment and cost estimate for the repair) is carried out by the Scottish Tractable company within 30 seconds. The “Allianz Schaden Express App” in Austria also automatically decides on cases but usually with a human in the loop[11].

Example: Behaviour-based telematic insurance premiums in life and car insurance

With the increasing spread of telematics insurance schemes in car insurance but also life insurance (e.g. the Vitality programme at Generali[12]), a new quality in the structure of insurance relationships was reached. For the first time, data about the individual behaviour of consumers is monitored and included in the pricing of net insurance premiums. This involves decisions about the consumer, which are made automatically by the system, as well as decisions in which a person (formally) takes the final decision but where AI systems play a decisive role in the preparation of that decision (for instance by profiling/categorising consumers). That telematics have the potential to fundamentally change the insurance sector is highlighted by the findings of a survey by the government-appointed German Advisory Council for Consumer Affairs in which some providers indicated “that the use of telematics will radically transform business models in the industry and that the trend towards behavioural tariffs is irreversible.”[13]

To account for the cases of high-risk applications outside the predefined sectors, the White Paper acknowledges, that there may be “**AI applications for certain purposes”** outside these sectors that should be considered **high-risk** and for which the legal requirements should apply equally. Given the need to address employment equality, the White Paper cites in particular that “AI applications for recruitment processes as well as in situations impacting workers’ rights would always be considered “high-risk””. Furthermore, the White Paper says that specific applications affecting consumer rights could be considered high-risk.

[11] *FAZ*, ‘*Der vollautomatische Kfz-Sachverständige*’, 2018 <https://www.faz.net/aktuell/finanzen/meine-finanzen/versichern-und-schuetzen/kuenstliche-intelligenz-in-der-kfz-versicherung-eine-revolution-15374987.html>.

[12] *Generali Vitality GmbH*, ‘*Generali Vitality*’, 2020 <https://www.generalivitality.com>.

[13] *Advisory Council for Consumer Affairs*, *Consumer-Friendly Scoring. Report of the Advisory Council for Consumer Affairs* (Berlin, 2018) <https://www.svr-verbraucherfragen.de/wp-content/uploads/Report.pdf>. p.71

> vzbv welcomes that the European Commission explicitly acknowledges the relevance of these applications affecting consumer rights. Still, vzbv considers this approach methodologically inconsistent. Therefore, vzbv proposes to supersede the current cumulated approach with a horizontal approach that will be complemented with sector specific rules at a later stage. As demonstrated, there are several good reasons for the European Commission to pursue a horizontal approach:

Increase methodological simplicity and consistency of future AI regulation

With the current cumulative and narrow approach, relevant authorities need three or more methodological frameworks of criteria to identify high-risk-applications:

1. Set of criteria to identify high-risk-sectors
2. Set of criteria to identify high-risk AI applications within these high-risk sectors
3. Set of criteria to identify high-risk AI applications outside these high-risk sectors

The third set of criteria to identify high-risk AI applications outside these high-risk sectors must – by definition -be universally applicable to all sectors, including high-risk applications within the high-risk sectors listed in the White Paper. Therefore, it is redundant to define two additional sets of criteria for the cumulative approach to identify high-risk sectors and high-risk AI applications within the sectors. In addition, to apply various criteria to identify high-risk applications is always prone to inconsistency.

The principle of regulatory simplicity imposes that the European Commission simply sticks to one set of criteria to identify high-risk applications across all sectors.

Increase legal certainty with a horizontal, gradual approach

The current proposal includes the option for policy-makers to define “AI applications for certain purposes” that are high-risk in addition to those within the predefined sectors (e.g. applications impacting workers’ rights, consumers’ rights). This approach creates a high degree of legal uncertainty for stakeholders. For providers and developers of AI it is not clear in how far their AI application will be considered high-risk or not. The same holds for consumers and the competent authorities. To define high-risk AI applications across all sectors with a single set of criteria increases legal certainty for all stakeholders: It offers clear guidance to all stakeholders as to what kind of AI applications are considered high-risk and which are not.

The German Data Ethics Commission made a proposal for a horizontal piece of regulation complemented by sector-specific rules. That proposal should be considered by the European Commission and the method for defining the risk-level of AI applications should follow this horizontal-sectoral approach accordingly:

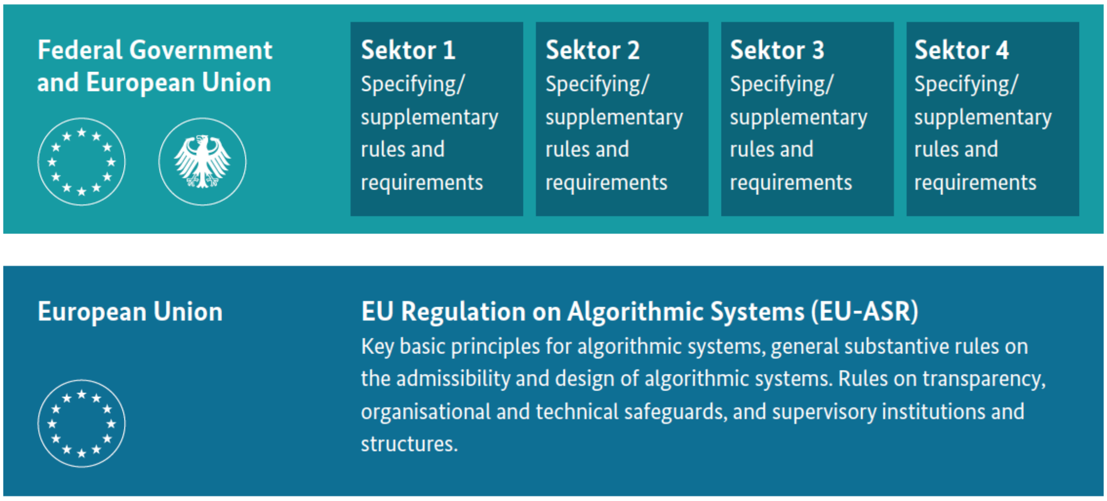
Federal Government and European Union
Sektor 1 Specifying/ supplementary rules and requirements
Sektor 2 Specifying/ supplementary rules and requirements
Sektor 3 Specifying/ supplementary rules and requirements
Sektor 4 Specifying/ supplementary rules and requirements
European Union
EU Regulation on Algorithmic Systems (EU-ASR)
Key basic principles for algorithmic systems, general substantive rules on the admissibility and design of algorithmic systems. Rules on transparency, organisational and technical safeguards, and supervisory institutions and structures.

Figure 1: EU regulation on algorithmic systems enshrining horizontal requirements and specified in sectoral instruments. Source: Data Ethics Commission (2019) Opinion of the Data Ethics Commission, p. 180, https://www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_EN_node.html

> vzbv proposes to replace the currently envisaged cumulated approach to risk-mitigation with a horizontal approach that will be complemented with sector-specific rules at a later stage. This ensures that all high-risk applications will be covered. It allows the EU to pursue a much more straight forward regulatory approach: It is methodically simple, increases plannability and legal certainty for companies and providers of AI, consumers, competent authorities and alike.

2.4 A more gradual and granular risk based regulatory approach

The European Commission rightly develops a vision to subject AI applications to a risk-based legal regime. Unfortunately, the White Paper proposes a binary approach: Applications are either high-risk or not. vzbv as well as the German Data Ethics Commission, business and the academic community in general all argue in favour of a more granular and gradual regulatory regime. The German Data Ethics Commission, for example, recommends classifying applications into five risk categories, each subject to a higher degree of regulation.[14]

The reasons in favour of a more granular risk-based approach than a binary one are various: The level of risk and the nature of the potential damage are different from application to application. But a granular approach can take this into account and subject each AI application to exactly the level of regulatory intervention that is appropriate for the risk it poses.

A binary "one size fits all" approach is to coarse and cannot deliver such a fine-tuned regulation. It must – and that seems unjustified - impose the same regulatory requirements on all AI applications within its scope. This means for example that the same requirements apply for applications that merely pose some potential for harm (or some

[14] Data Ethics Commission. P. 180; Martini.; *Katharina Zweig and Tobias Krafft*, '*Transparenz und Nachvollziehbarkeit algorithmenbasierter Entscheidungsprozesse - Ein Regulierungsvorschlag aus sozioinformatischer Perspektive*', 2019 <https://www.vzbv.de/pressemitteilung/algorithmen-kontrollieren-geltendes-recht-durchsetzen>.*AI Ethics Impact Group*, *From Principles to Practice - An Interdisciplinary Framework to Operationalise AI Ethics*, 2020 <https://doi.org/10.11586/2020013>. See also *AI Ethics Impact Group*, *From Principles to Practice - An Interdisciplinary Framework to Operationalise AI Ethics*, 2020 <https://www.bertelsmann-stiftung.de/fileadmin/files/BSt/Publikationen/GrauePublikationen/WKIO_2020_final.pdf>.

elevated level of risk) as on those applications with a serious potential for harm (very high level of risk). This means that the requirements will be too harsh on some applications and too lenient on others. As a consequence, a binary approach must exclude applications below a certain relatively high threshold, from all regulatory requirements to not be disproportionate, leavening consumers and society exposed to the risks they pose.

Both regulatory downsides of the binary approach are not satisfactory from the policymakers point of view. Therefore, the European Commission should adopt the granular risk-based regulatory approach recommended by the German Data Ethics Commission (compare Figure 2).

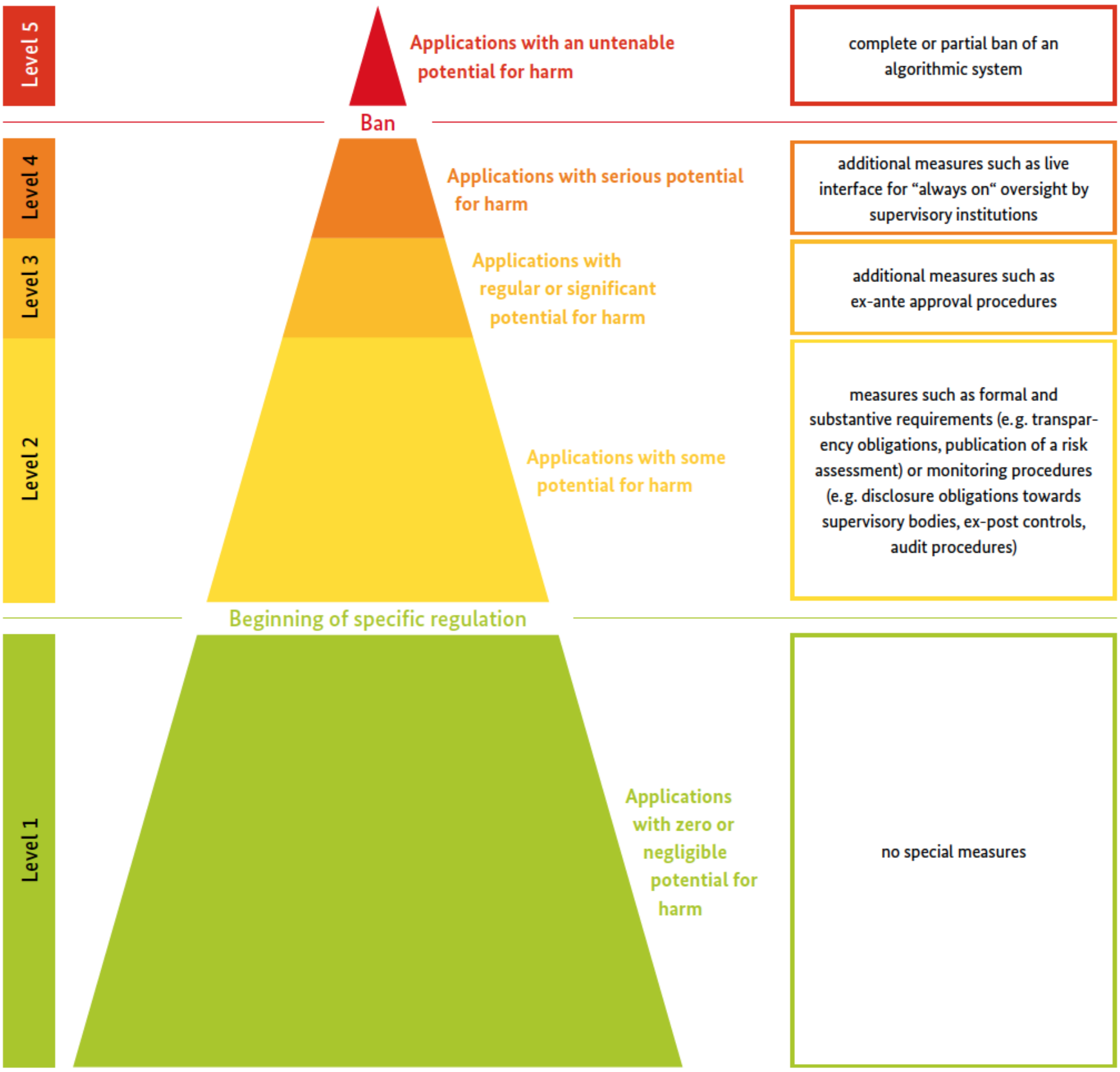
Level 5
Applications with an untenable potential for harm
complete or partial ban of an algorithmic system
Ban
Level 4
Applications with serious potential for harm
additional measures such as live interface for "always on" oversight by supervisory institutions
Level 3
Applications with regular or significant potential for harm
additional measures such as ex-ante approval procedures
Level 2
Applications with some potential for harm
measures such as formal and substantive requirements (e.g. transparency obligations, publication of a risk assessment) or monitoring procedures (e.g. disclosure obligations towards supervisory bodies, ex-post controls, audit procedures)
Beginning of specific regulation
Level 1
Applications with zero or negligible potential for harm
no special measures

Figure 2: Criticality pyramid and risk-adapted regulatory system for the use of algorithmic systems. Source: Data Ethics Commission (2019) Opinion of the Data Ethics Commission, p. 180, https://www.bmjv.de/DE/Themen/FokusThemen/Datenethikkommission/Datenethikkommission_EN_node.html

> Risk and corresponding risk-mitigating legal obligations should be assigned gradually. To simplify, similar risks and corresponding obligations could be grouped into a number of categories. The obligations should reflect the corresponding risk level:

Applications posing a higher risk should be subject to stricter, more demanding mandatory requirements. The granular approach ensures that all risky applications across all sectors are subject to adequate, application-specific mandatory requirements, thereby reducing the regulatory burden and increasing legal certainty for all stakeholders.

2.5 The definition of risk and damage: Include economic and monetary damages

Many individuals of a specific group may systematically suffer small losses (e.g. through price differentiation, or denial of certain service levels). The individual economic loss/damage may be small, but the aggregated effects can sum up to considerable sums.

The assessment of the level of risk of a given AI application must be based on all potential impacts on the affected parties. For instance, applications that produce legal or similarly significant effects for the rights of an individual or a company; that pose risk of injury, death or significant material or immaterial damage; that produces effects that cannot reasonably be avoided by individuals or legal entities.

The definition of damages and harm must also explicitly include monetary and economic losses and access to markets/scarce resources (e.g. when platforms deny access to markets (like AirBnB) or poor credit conditions).

3 MORE TRANSPARENCY FOR CONSUMERS

vzbv appreciates that the White Paper proposes a wide range of mandatory requirements and outlines some features of a potential ”conformity assessment” for high-risk AI applications.

The proposed mandatory transparency requirements towards consumers merely include the labelling of high-risk AI applications so that users know that they interact with a machine. Although vzbv holds that labelling is important, it must be stated that more requirements generating transparency towards consumers are needed: consumers need to know about the risks and the reliability of an AI application, the data that is underpinning the decision and how a specific decision came about. Developers and operators of AI systems must be able to explain how their systems work to ensure traceability (and accountability). Data subjects as defined in GDPR must be provided with all the information necessary to exercise their rights when necessary.

vzbv suggests adopting a more granular approach where legal obligations are applied more gradually, reflecting the risk level of the specific AI application. Overall, vzbv regrets that the White Paper proposes too few legal requirements with respect to transparency for consumers. Therefore, vzbv suggest some modifications of those requirements.

3.1 Information provision: publication of a risk assessment

In section c) of the White Paper, the European Commission proposes to make it obligatory to provide developers with information about an AI system’s capabilities and limitations (including its level of accuracy). The White Paper rightly states that such information could “be relevant to competent authorities and affected parties”. vzbv is under the impression, that the European Commission conceived this information provision in order to provide competent authorities and “affected parties” with a risk assessment of the AI system. However, vzbv recommends the modification of the concept of the risk assessment in the White Paper:

Addressees of the risk assessment must include consumers and the public

Trust in AI can only emerge on the basis of an informed public debate and assessment of the risks and opportunities of these systems. It is vital for consumers to be included in such a public debate. Therefore, the European Commission should clearly state that consumers/citizens are considered “affected parties” under this regulatory framework.

In addition to individual consumers and interested members of the public, such a debate naturally must include policy-makers, scientists, civil society organisations and affected businesses. Furthermore: In order to enable such a debate, the basic characteristics and potential risks of AI systems must be made publicly available. For example, operators of AI applications must provide Information on what general data bases a machine learning system has been trained, which data categories are used to make/prepare decisions, according to which criteria the system optimises, how accurate a system is and how bias has been reduced.

Clearly, the information entailed in the risk assessment must not include business secrets. Nonetheless, the information must be specific enough to allow for an informed debate.

Comprehensive risk analysis: enabling public debate and trust

In order to enable a public debate, the European Commission should clarify that it will demand a comprehensive risk analysis that also considers systemic risks in a wider socio-economic context: therefore, the obligation to publish a risk assessment should also apply for the processing of non-personal data.

The published risk assessment should take into account considerations that lie outside the realm of data protection: risks to self-determination, to privacy, to physical integrity, to personal integrity and to property as well as a risk of discrimination. In addition, the risk impact assessment should take into account not only the underlying data and the logic of the model, but also measures of quality and fairness of the underlying data models, e.g. bias (distortions) or (statistical) error rates that a system exhibits when making predictions or forming categories.

> vzbv suggests that the European Commission explicitly clarifies that the notion of “affected parties” includes consumers.
>
> vzbv furthermore suggests that the European Commission obliges operators of AI applications to publish a risk impact assessment.
>
> The European Commission should clarify that such a published risk assessment must not contain business secrets but information that the public needs to conduct an informed debate. The risk assessment should take into account a comprehensive set of risks, including among others: risks to self-determination, to privacy, to physical integrity, to personal integrity and to property as well as the risk of discrimination. It should entail comprehensive information about the underlying data fed into the algorithm and the logic of the model, robustness, accuracy and fairness of the underlying data and models.

3.2 Individual explanation for affected persons

The White Paper proposes a number of obligations for high-risk algorithmic systems (including transparency vis-à-vis supervisory authorities, documentation requirements and a labelling requirement). Regrettably, there is no provision obliging operators of the systems to provide an individual explanation of the specific reasons for a decision to those affected.

Such information rights are central for consumers to be able to understand and individually review algorithm-based decisions. Only then can they exercise their rights - as laid down in the GDPR, for example - and challenge a decision on a well-founded basis. For example, to defend themselves against discrimination or erroneous decisions.

> vzbv proposes, that the Commission amends the Whitepaper with a provision mandating providers of high-risk AI applications to inform consumers and explain the result of the individual case in a comprehensible, relevant and concrete manner (in contrast to the general duty to inform under the GDPR, where the functioning of an AI or algorithmic system is explained in general terms).

4. KEEPING OF RECORDS AND DATA

vzbv welcomes that the White Paper proposes various obligations to retain data and records. The European Commission – and rightly so – pursues the objective to enable competent authorities to perform controls and checks of AI systems. This requires that decisions of AI systems should be traceable and verifiable. The obligation to retain and document data includes information on training agendas, training methods, methods for programming, training, testing and validation of the systems.

> vzbv regrets that the European Commission notes that datasets themselves should only be kept in certain cases, without specifying the criteria for identifying these cases. A granular approach to classify applications according to risk as proposed by vzbv would allow the regulator to define a risk-level above which the datasets of AI systems must be retained for later inspection by the authorities.

5. SPECIFIC REQUIREMENTS FOR REMOTE BIOMETRIC ANALYSIS

vzbv welcomes that the European Commission devotes a separate paragraph to the topic of remote biometric identification. However, vzbv recommends that the European Commission considers several amendments to this topic:

5.1 The scope should include biometric analysis and metabolites

vzbv recognises the importance of the topic of remote biometric identification yet emphasises that remote biometric identification covers only one spectrum of (remote) biometric analysis which can go far beyond the mere remote identification of persons and is highly sensitive. Biometric analysis includes among others the analysis of current moods and personality traits of people based e.g. on the analysis of speech and facial expressions going as far as lie detectors based on the analysis of eye movements and changes in pupil size[15].

[15] *The Guardian*, '*The race to create a perfect lie detector – and the dangers of succeeding*', 2019 <https://www.theguardian.com/technology/2019/sep/05/the-race-to-create-a-perfect-lie-detector-and-the-dangers-of-succeeding>.

Remote biometric analysis of consumers is particularly sensitive. It can constitute an intrusion into the personality rights of consumers. Also, biometric analysis of consumers by businesses can greatly increase the prevalent structural imbalance between consumers and companies: Private companies using biometric analysis will have many more pieces of personal and sensitive information at their hands when "negotiating" with consumers. One example is real-time video-based analysis of a customer's behaviour in a supermarket in order to personalise offerings.[16]

In addition to the much-noticed biometric identification and analysis, the collection and analysis of so-called "metabolites"[17] can be particularly risky. This involves the analysis of particles that people leave behind/give off, for example sweat, dust, breath, etc. The analysis of metabolites can allow conclusions to be drawn about individual behaviour, consumption and habits and is thus highly sensitive. Therefore, it should be included in the scope of the regulation and be subject to strict legal limits.

> vzbv recommends that the European Commission broadens the scope of a future regulatory framework by including remote biometric analysis in general.
>
> Furthermore, vzbv proposes that the scope of the future regulatory framework also explicitly includes the analysis of so-called "metabolites".

5.2 Distinguish between public and private spaces as well as public and private uses

vzbv recommends that the European Commission clearly distinguishes between the use of remote analysis of biometric/metabolite information between public and private spaces as well as public and private uses. In general, all forms of biometric/metabolite analysis of consumers should be considered highly risky and therefore subject to the strictest regulatory requirements.

> The operation of remote biometric identification systems by public and non-public operators should be prohibited in public places (including publicly accessible places such as supermarkets) until the associated risks and consequences for individuals and society have been adequately researched. The same should apply for the analysis of metabolites.

The European Commission explains that biometric remote identification is already regulated by the GDPR. As outlined above, given the role of profiling and preparation of human decisions by AI, which are not sufficiently covered by the GDPR, the GDPR cannot provide sufficient protection for consumers.

6. GOVERNENCE

The White Paper proposes a European governance structure for AI. However, the proposed governance structure should perform rather "passive" tasks, like support the coordination and cooperation of national and European supervisory authorities, support standardisation etc.

[16] *Innovative Retail Laboratory*, '*VICAR*', 2018 <https://www.innovative-retail.de/news/vicar.html; https://www.nec.com/en/global/techrep/journal/g18/n02/180210.html>.

[17] *The Economist*, '*People leave molecular wakes that may give away their secrets*', Feb 13th 2 <https://www.economist.com/science-and-technology/2020/02/13/people-leave-molecular-wakes-that-may-give-away-their-secrets>.

A thorough audit of an AI system, performed by a component authority, is a highly complex task. It requires, among other things, a high level of technological and methodological expertise (e.g. in data science, statistical analysis etc.). This expertise cannot be taken for granted.

Therefore, vzbv regrets that the White Paper does not lay out plans on how a European governance structure could actively support the respective competent authorities in their daily supervisory activities, e.g. with checking whether the obligatory conformity assessment for high-risk applications had been carried out in a diligent manner. The German Data Ethics Commission calls for the creation of "national competence centres" that support the existing sector-specific supervisory authorities: with methodological and technical expertise, in particular, lending support for the competent authorities with checking the extent to which algorithmic systems comply with the law.

> The Commission should adopt this approach and work towards a European governance structure or agency that supports the sector-specific national and EU competent authorities directly in supervising AI systems with methodological and technical expertise. This could be particularly helpful for competent authorities in smaller Member States who might find it difficult to build up such technical and methodological competence themselves.

7. INTERNATIONAL TRADE

Unless and until the European Union does not dispose of a functioning regulatory and governance structure for AI, the European Commission should put extra care on the consistency of its overall policies in the field of AI. The EU needs to ensure that the European approach to AI is not compromised by international trade commitments that might limit the EU's regulatory autonomy in the long run. The Joint Statement Initiative on e-Commerce under the roof of the World Trade Organisation (WTO) requires specific attention in this regard.

> The European Commission needs to ensure that its approach to the regulation of AI is consistent throughout the EU's policy areas especially regarding commitments in international trade.

8. LIABILITY

vzbv welcomes that the White Paper points out that AI poses new risks for an effectively functioning liability regime. Consumers that incur damage from the intended use of AI-powered products and services must have effective and easy-to-use tools to claim compensation from the responsible parties. The existing EU liability regime does not provide these tools.

When a consumer purchases AI-powered products and services, the vendor or deployer can often absolve themselves for any damage caused by occurring faults. The vendor or deployer will in these cases claim to have been uninvolved in the creation process and hence will not assume responsibility for software faults, instead referring to the developer or another third party. Accordingly, vzbv holds that it is mainly product liability law, targeted at the manufacturer of products, that can help consumers reach effective compensation. To be more accurate with regard to AI-powered products and services, vzbv will be referring to the product creator as "developer" instead of "manufacturer". In some cases, AI-powered products and services cannot be attributed to a single developer but are instead deployed by a third party. For example, when online platforms provide an AI-powered service to the consumer, it can be difficult to ascertain the exact creator of different parts of the service. Accordingly, vzbv holds that it is important to hold the deployer of AI-powered products and services liable as well.

The existing body of EU product liability law is dominated by the EU Product Liability Directive 85/374/EEC (PLD) which was designed without AI-powered products and services in mind, not mentioning AI-powered products. As a result, it falls short of the challenges posed and the damage caused by increasingly technically complex and connected device. To implement an effective and just liability regime at European level, vzbv believes that the European Commission should substantially revise the PLD.

8.1 Update the general scope of the Product Liability Directive

While it is widely believed that software is a product in the sense of the PLD, the definition of the term “product” does not explicitly clarify whether that is actually the case. The European legislator should adjust the PLD’s definition of “product” to provide clarity.

> The European Commission should revise the PLD and propose a scope that includes all forms of technologies and objects. The Directive should cover devices as well as their function and effect. This includes software and digital content no matter whether it is embedded in a physical product or not.

8.2 Reconsider the term “defect”

The opacity and complexity of AI-powered digital products, paired with crucial information often being purposefully hidden as business secret, make it extremely difficult for consumers to prove

- that there is a defect in the product in the first place,
- that damage has occurred and
- the causal link between the two.

That already holds true for cases where a consumer is damaged by a single, easily identifiable product. In case of networks of multiple AI-powered products and services, such as an AI-based smart home system, it can be impossible for the consumer to even identify the defective product. When different manufacturers and/or operators provide collective networks to the consumer, each can refer to the others in case of a defect. The consumer, on the other hand, has no access to the data necessary to identify the defect’s source.

To remedy this situation, the first step would be to reconsider using the term “defect” altogether. It assumes a product over which, once sold, the manufacturer has no more influence. Under this assumption, it makes sense to restrict liability according to the state of the art and available scientific and technical knowledge at the time of placing the product on the market (Article 7 (e) PLD).

However, this does not reflect the reality of the market for AI-powered products and services. AI-powered products remain connected to the developer long after sale. The developer or deployer is often the only one who fully understands the product’s functionality. In the case of AI-based products or services with learning components the connection to the developer or deployer over the whole life-cycle is profound: Developers or deployers of AI-based products or applications regularly adapt the systems after the consumer has acquired or subscribed to it. Adaptations can occur whenever the developer or deployer (or another party) applies updates to a digital product or based on machine learning techniques. This can alter the characteristics of the system significantly, e.g. by changing the weighs of decision criteria or including new (types of) data in the decision model.

Similarly, security can usually only be maintained by the developer or deployer through periodic software updates.

Accordingly, the developer and the deployer must assume responsibility for the functioning and the effects of their AI-powered products and services, independently of a defect that can be proven.

> The European Commission, in a review of the PLD, should propose a product liability regime that is independent from a defect: Developers and deployers should bear strict liability whenever a product causes damage while used as intended.

8.3 Adjust the rules for the burden of proof

Since the consumer has no access to the relevant data, the parties who do should bear the burden of proof. Whoever creates an AI-powered product or service must, if at all possible, enable the logging and processing of all data needed to identify who is liable in case of damage. If logging and processing are for some reason impossible, this should be to the detriment of the developers or deployers who derive commercial gain from the product or service. It should make no difference whether that party is the manufacturer in the sense of the PLD, or involved in some other form, e.g. as developer, deployer, operator or service provider.

The European Commission should therefore propose strict manufacturer liability as well as a joint liability regime whenever several developers or deployers participate in a single AI ecosystem. When designing such a joint liability regime, the European Commission should also consider all other involved actors, such as operators or service providers.

> In ecosystems of connected AI-powered products and services, all developers and deployers involved in providing the product to the consumer should bear joint liability.

8.4 Broaden the PLD’s scope with regard to types of damage

In addition, vzbv holds that the type of damages covered by the PLD are outdated as it only covers damages such as death, personal injury or damage to property (Article 9).

At the same time, information, data and software are of ever-growing relevance not only to the economy as a whole but also to consumers. Damage done by AI-powered products and services is often intangible but nonetheless severe. A software deleting or leaking important consumer data causes damage that is just as real as the software itself. The PLD must reflect these developments to be relevant in the digital age.

> The European Commission should propose a new definition of damage that also covers immaterial damage to consumer data or their digital environment as well as leakage of consumer data.

In addition, the European Commission should consider widening the definition of damage to include more forms of damage, e.g. to include financial and economic losses.