ohne Starke Verbraucher keine Starke Wirtschaft

verbraucherzentrale Bundesverband

ARTIFICIAL INTELLIGENCE NEEDS REAL WORLD REGULATION

Position paper of the Federation of German Consumer Organisations (vzbv) on the European Commission's proposal for an Artificial Intelligence Act (AIA)

05 August 2021

Impressum

Verbraucherzentrale
Bundesverband e.V.

Team
Digital and Media

Rudi-Dutschke-Straße 17
10969 Berlin

digitales@vzbv.de

Bundesverband der Verbraucherzentralen und Verbraucherverbände

Verbraucherzentrale Bundesverband e.V.

CONTENTS

I. SUMMARY

vzbv welcomes that the European Commission proposes a regulation laying down harmonised rules on artificial intelligence (AI) in the (Artificial Intelligence Act (AIA)). The AIA must mitigate AI-related risks for consumers. To achieve this goal and accomplish the European Commission's stated objective of a 'trusted AI', vzbv recommends to increase the focus on consumers of the AIA and to strengthen the possibilities for independent assessments of high-risk AI systems.

The broad definition of AI systems ensures the AIA is future proof

vzbv welcomes the broad definition for AI systems. It covers the relevant a range of algorithm-based decision making Systems (ADM) and the underlying many AI systems that prepare or make vital decisions on consumers. It also corresponds to the definition of ADM used in the scientific community[1]. The broad definition of AI ensures that the AIA will be future proof. Restricting AI-rules on a narrow set of AI, like machine learning, risks outdating the AIA soon, when new AI techniques emerge.

Consumers must receive individualised explanations

AI systems must be transparent and comprehensible in order to enable sovereign consumer decisions. Regrettably, the draft AIA provides no transparency towards consumers beyond a labelling obligation for three types of AI systems (Art. 52).

The AIA must contain a provision mandating **providers** or **users of high-risk AI systems** to inform consumers and **explain** the **result** of the individual case in a comprehensible, relevant and concrete manner (upon their request). Such information rights are central for consumers to be able to understand and individually review an AI system's decision. Only then can consumers can exercise their rights. This must include information on the **input data** on the basis of which an AI application made/prepared a decision about the individual, the **logic of the model**, the **criteria** against which the AI system optimises, **measures** of **fairness**/bias, **robustness**, and **accuracy** as well as the **purpose** of the use of the AI system.

The general public needs information

Trust in AI systems can only emerge on the basis of an informed public debate and an assessment of the risks and opportunities of these systems. Providers of high-risk AI systems must provide the **public** with meaningful **information** that is relevant for an **informed debate** and understanding of an AI system. It should include for example information on the characteristics, capabilities and limitations of performance of the system as well as information on human oversight, corresponding to the information obligations towards professional users (Art. 13).

Complement the list of high-risk AI systems

The current AIA proposal focuses on AI-risks related to (product-)safety, health and fundamental rights linked to the use of AI systems. It thereby neglects that AI systems can cause significant economic/financial welfare losses for consumers or lead to violations of consumer rights.

---

[1] 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>.

AI systems operating in the following areas must be **included** in the list of **high-risk applications** in Annex III as they can cause serious economic/financial harm to consumers or severely violate consumer rights: AI systems intended to be used in the area of **insurances**, consumer-facing AI applications for **financial investment** or **portfolio management**, payment and debt collection. Also for **scoring and profiling** of consumers when they determine consumers' access to services or markets and AI systems determining consumers' access to the **housing market** should count as high-risk.

Future high-risk applications: Consider consumers rights and economic risks

To make the AIA **future proof**, **legislators** must be able to **add new high-risk AI systems from other areas of application** than those already listed in Annex III**.** When updating the list of high-risk AI applications in Annex III legislators can only add AI systems that "pose a risk of harm to the health and safety, or a risk of adverse impact on fundamental rights" (Art. 7 par. 1. (b)). To ensure the AIA is future proof, **legislators** must be able to **declare** an **AI systems as high-risk** when it poses significant risks with respect to the **violation of consumer rights**, as well as **social** and **economic harms** for individuals and (social) groups.

Prohibit dark patterns and exploitation of consumers' vulnerabilities

Art. 5 par. 1 (a) should in general **prohibit** that AI systems exploit so-called **"dark patterns"** by presenting end user choices in a non-neutral manner, or by otherwise subverting or impairing user autonomy, decision-making, or choice via the structure, function or manner of operation of a user interface or a part thereof. Also Art. 5 par. 1 (b) should in general **prohibit** AI systems from **exploiting vulnerabilities** of **consumers**. It should not be limited to young, old and persons with disabilities, but include all consumers. Every person can find itself in very vulnerable positions temporarily. AI systems must not exploit vulnerabilities caused by emotional distress, exhaustion, tiredness, grief, sorrow, physical pain or the influence of medication.

The AIA should **not require intentionality** as a precondition for prohibiting dark patterns or for **exploiting** weaknesses and **vulnerabilities** of **consumers**. It is near to impossible to prove that providers of AI systems intended harm. Also, the harm caused by these systems should **not** be limited to physical or psychological harm, but also **include socio-economic welfare losses**, violations of **fundamental rights** (e.g. discrimination) and **consumer rights** (e.g. deception).

Prohibit general social scoring by private entities

Art. 5 par. 1 (c) bans general scoring by public authorities under certain circumstances. General social scoring undertaken by private entities can also have a large negative impact on individuals or entire groups. It can lead to unjustified exclusion of consumers from entire markets or services, discrimination, economic and financial harm. The AIA must **prohibit** that **private entities** use **social scoring** for the evaluation or classification of people's trustworthiness based on their social behaviour or to predict personal or personality characteristics.

Prohibit remote biometric identification in public spaces by private entities

The use of biometric identification systems in publicly accessible spaces can cause significant harm to consumers, including severe violations of the right to privacy and of their autonomy. Examples of potentially harmful applications include smart glasses, augmented reality applications on mobile phones or analysis of shopping centres' surveillance camera footage. The AIA must **prohibit** the use of **biometric identification**

systems in publicly accessible spaces by **private entities** (not only public authorities). The ban must include 'real-time' as well as retrospective biometric identification.

Prohibit emotion recognition system by private entities

The AIA does not protect consumers effectively from private entities using emotion recognition systems. Companies can exploit these to deceive and manipulate users, to undermine or subvert consumer autonomy and decision-making, using automated recognition and analysis of human features, facial expressions, voice, keystrokes and other biometric or behavioural signals. Art. 5 must ban the use of **AI-based emotion recognition systems** and the analysis of consumers' emotions by **private entities**. **Exception** from the ban should be granted for specifically defined purposes that are to the clear and proven **benefit of the consumer** (such as for medical or research purposes in the public interest) in strict compliance with applicable data protection law and subject to appropriate safeguards.

vzbv joins the EDPB's demand for a "general **ban** on any use of AI for **automated recognition of human features** in publicly accessible spaces".

Independent conformity assessment for all high-risk AI systems

Legislators must not leave the conformity assessment of high-risk AI systems to the AI providers' self-assessment. Only independent audits can create consumer's trust and foster the acceptance of AI in general. **All high-risk** applications must be subject to **independently** verified **conformity assessments** as laid out in Annex VII when the AI system is brought to **market for the first time**. Also the AIA should demand independent conformity assessment in case of well funded **indications** that high-risk AI systems are **not in conformity** with the AIA's requirements (e.g. when the system has been changed, or is employed in another context).

Complement enforcement with independent assessments on request of civil society organisations

Limiting **market surveillance** to public **authorities** and institutions (Art. 64 par. 3) is **not sufficient**. **Civil society organisations** must have the right to **request audits** of high-risk AI systems by notified bodies when there are reasonable indications that the high-risk AI system violates European or Member States' legislation, has a significant negative impact on the social, economic, physical or psychological wellbeing or the security of persons or social groups, or poses significant environmental risks.

Legislators must establish **due process obligations** for providers of high-risk AI systems so that notified bodies, on the request of civil society organisations, can conduct independent audits. This must include obligations for providers to give auditors access to all data, documentation and records needed to assess the AI systems' risks to the social, economic, physical or psychological wellbeing and security of persons or groups as well as its potential environmental impact. The notified body must publish the findings of the **audit** in a **report**.

Ensure private enforcement of the AIA

Consumers greatly benefit when consumer organizations enforce their rights complementary to enforcement by competent authorities. To **ensure** that **consumer organizations** can **enforce** the **AIA** provision in **courts** legislators must add the AIA to Annex I of the European Directive on representative actions for the protection of the collective interests of consumers ((EU) 2020/1828).

Trade agreements must not hinder an effective transparency and monitoring of AI systems

Current EU trade negotiations might restrict the Europe Union’s ability to regulate in the field of AI in the future, in particular with regard to independent assessments and audits. Legislators must enact **trade rules** that do **not impede** on future **AIA rules**.

II. INTRODUCTION

This statement provides the Federation of German Consumer Organisations' (Verbraucherzentrale Bundesverband - vzbv) feedback to the European Commission's proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act (AIA)).[2]

vzbv welcomes the opportunity to comment on the European Commission's proposal, as artificial intelligence (AI) increasingly shapes consumer markets and our societies. AI systems in consumer markets undoubtedly benefit consumers in some areas. This is for instance the case when improving personalisation of services and thus increasing convenience for instance in e-commerce, individualising health care, automation of vehicles and providing driver support systems in cars. On the other hand, obscure AI also increases risks for consumers, for example, by enabling undertakings to exploit personal vulnerabilities or infere with consumer preferences, by enabling discriminatory practices, unfair treatment or violations of people's privacy. The AIA must strike a balance to mitigate these AI-related risks so consumers can reap its benefits. To achieve this goal and accomplish the European Commission's stated objective of a 'trusted AI', vzbv recommends to increase the focus on consumers of the AIA and to strengthen the possibilities for independent assessments of high-risk AI systems.

[2] European Commission: Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021) 206 final) (hereafter 'AIA') (2021), URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206 [Access: 20.07.2021].

III. PROPOSALS FOR INCREASING THE AIA'S CONSUMER FOCUS

1. DEFINITIONS

1.1 Article 23 (1) – Artificial intelligence system

vzbv welcomes the broad definition for 'artificial intelligence system' (AI system) in Article 3 (1) and the list in Annex I. Annex I lists a range of algorithm-based decision making Systems (ADM) and their underlying techniques. These are embedded in a wide range of AI-driven applications where they prepare or take decisions about consumers or on their behalf. These in turn can have a substantial (negative) impact on peoples' lives in various contexts.

The definition of AI systems as provided in Article 3 (1) in conjunction with the list of approaches and techniques in Annex I is appropriate, and not too broad. It corresponds to the definition of ADM used in the scientific community[3] as systems that prepare or even make decisions over the treatment of consumers. It is also in line with the recommendations of the Data Ethics Commission for the German Federal Government[4] which refers to "algorithmic systems". The actual goal of the AIA is regulating and protecting consumers from harms caused by such systems.

The broad definition of AI ensures that the AIA will be future proof. Restricting AI-rules on a narrow set of machine learning systems misses the point. A narrow definition of AI systems risks outdating the AIA soon, when new AI techniques emerge, as new forms of fast development of AI techniques in the past have demonstrated.[5]

> European legislators should **keep the broad definition of AI systems** laid out in Article 3 (1) and Annex I. They reflect algorithm-based decision-making systems that underlie many critical AI systems that prepare or make vital decisions on consumers.

1.2 Article 3 (4) – User

Article 3 (4) defines the 'user' of an AI system as professional users only. Art. 3 lacks a definition for the non-professional user, e.g. people using AI driven health application giving them health advice[6], or consumers affected by AI-driven decisions of systems employed by professional users. The omission illustrates the lack of consumer focus that runs throughout the whole AIA draft.

---

[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>.

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

[5] For example, see the tremendous advancements of neural network-based AI systems over the past decade: Pitchford, Daniel. Forbes: A Decade Of Advancements As We Enter A New Age Of AI (2019), URL: https://www.forbes.com/sites/danielpitchford/2020/12/31/a-decade-of-advancements-as-we-enter-a-new-age-of-ai/?sh=5c65d3cb4055 [Access: 19.07.2021].

[6] Ada Health GmbH: Ada Website, URL: https://ada.com/app/ [Access: 29.06.2021].

Legislators must complement Article 3 with a **definition** for **non-professional users** of AI systems. This must entail people using AI systems in their capacity as **consumers** and **citizens**. It must also consider **consumers** who are **affected** by AI systems employed by **professional users**.

1.3 Article 3 (34) – Emotion recognition system

The definition of 'emotion recognition systems' in Art. 3 (34) is too narrow. The definition relies on the definition of biometric data as defined in Art. 3 (33), which itself is taken over from the General Data Protection Regulation (GDPR).[7] It holds that biometric data must "allow or confirm the unique identification of that natural person." As a consequence 'emotion recognition systems' that do *not* rely on data allowing the unique identification of a natural person, will fall out of the scope of the AIA. However, vzbv holds that these types of systems should also fall under the AIA's scope.

This could include systems that rely only on the analysis of clicking, typing and cursor movement data for example. Also, for an AI system supporting a retail salesperson in a shop, it is not important to know the identity of a potential customer entering the shop. The AI system can provide the shop personnel with valuable real time personality/emotion analysis data, based on the customer behaviour. For example inferences from measures on the relative tone/height, rhythm, and the speed of a voice, but not the voice itself.

- The definition of 'emotion recognition systems' in Art. 3 (34) should not refer to biometric data but to personal data. Otherwise, there is a significant risk for circumvention of the legislation.

2. THE SCOPE IS TOO NARROW PART I: NEGLECT OF ECONOMIC HARMS AND VIOLATIONS OF CONSUMER RIGHTS

In general, the scope of the proposed AIA is too narrow and the legislation does not focus on consumers. The European Commission's proposal focuses on problems of (product-)safety, health and fundamental rights linked to the use of AI systems. It mostly deals with high risks to people in their capacity as citizens and employees, neglecting that AI systems can lead to significant economic/financial welfare losses for consumers or to violations of consumers' rights.

2.1 AI applications with large economic/financial impact or effects on consumer rights must be regarded as high-risk

The European Commission's proposal sees high-risks of AI systems nearly exclusively in the areas of (product-)safety, health and fundamental rights. The draft AIA focuses on mitigating risks to people in their capacity as citizens, patients, employees and students ("education"). However, most AI systems in these areas are already subject to European legislation. Therefore, in practices, it can be doubted that consumers will benefit much from the draft AIA in these areas.

[7] Compare Art. 4 (14): GDPR: European Parliament: EU General Data Protection Regulation (GDPR) Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, OJ L 119, 4.5.2016, 2016.

The proposed AIA widely neglects consumer rights outside these areas. In particular, AI systems employed by private entities that can lead to significant economic and financial welfare losses in consumer markets are out of the draft AIA's focus. Credit scoring is the only "classic" consumer market high-risk AI application listed in Annex III and typically provided by private companies.[8]

In vzbv's view, the AIA must include more areas of applications where AI systems are employed by private entities and can violate consumer rights or cause significant economic and financial welfare losses. Legislators must classify some applications that fall into these areas of application as high-risk systems. For others it will be sufficient to assign them to a to-be-defined medium risk category, with appropriate requirements and obligations. In this respect, vzbv regards Art. 52 as insufficient as a "medium-risk" category, as it merely entails labelling obligations for three selected areas of application.

The European Commission framed the AIA as a legislation fostering a "trusted AI". Unfortunately, neglecting a wide range of AI applications that can cause significant economic and financial harm for consumers or violations of consumer rights will further undermine peoples' low trust in this technology as recent surveys suggest.[9]

AI-systems used in the area of insurances are the most obvious example for AI systems that should be added to the list of high-risk AI systems. With the increasing spread of telematics insurance schemes in car insurance but also individual behaviour-based tariffs for life insurance (e.g. the Vitality programme at Generali), a new quality in the structure of insurance relationships is reached. For the first time, data about the individual behaviour of consumers is monitored and included in the pricing of net insurance premiums.[10] AI-driven individualised behavioural tariffs, e.g. in telematics-based motor insurance or in health insurance, are likely to become much more widespread and transform the entire insurance industry. Other areas of application include AI-based risk categorisation of consumers for individually determining insurance premiums for liability and household content insurances[11] and behaviour-based bonus calculations for

[8] Other areas AI-applications in in Annex III include infrastructure (Annex III 2.) or education (Annex III 3.). However, these are regularly provided by public institutions or highly regulated (e.g. emergency services).

[9] TÜV-Verband: Verbraucher wollen Sicherheit und Transparenz bei Künstlicher Intelligenz (2020), URL: https://www.tuev-verband.de/IG-NB/vdtuev-startseite/news/ki-studie?context=e3068ebc9b4940b0b56ad4576ca633bd [Access: 20.07.2021]; BEUC: Artificial Intelligence: what consumers say - Findings and policy recommendations of a multi-country survey on AI (2020), URL: https://www.beuc.eu/publications/survey-consumers-see-potential-artificial-intelligence-raise-serious-concerns/html [Access: 20.07.2021].

[10] Compare: Generali Vitality GmbH: Generali Vitality (2021), URL: https://www.generalivitality.com/about-our-program/ [Access: 20.07.2021].

[11] Taulli, Tom: Lemonade IPO Shows The Power Of AI (Artificial Intelligence) (2020), URL: https://www.forbes.com/sites/tomtaulli/2020/07/03/lemonade-ipo-shows-the-power-of-ai-artificial-intelligence/?sh=152fd0f83aeb [Access: 23.07.2021]; Lemonade: Lemonade Contents & Personal Liability Insurance | Protect The Stuff You Love, URL: https://www.lemonade.com/de/en [Access: 23.07.2021].,

life insurance[12], AI-based claims handling for car-[13], liability and household content insurances[14]. Insurers can already use individual behavioural-based bonus programs as vehicle to price particular consumer groups out of the market, thereby undermining the principle of solidarity.[15] Thereby AI can lead to significant unjustified treatment, discrimination and financial harm for individuals or groups of consumers.[16]

Another area in which AI systems can obviously lead to significant financial harm for consumers are consumer-facing AI applications intended to be used automated financial investment or portfolio management.[17]

AI-driven payment and debt collection services are another area of concern. The online shopping boom during the corona pandemic led to an increasing number of consumer complaints that revealed underlying problems of these services[18]: A large provider raised vzbv's attention as he rejects consumers' money transfers allegedly, because the stated purpose for the transaction does not exactly correspond to its specifications. The reason for the rejection is presumably a fully automated process. Consequently, the provider uses a debt collection agency to handle the case and charges the consumer for the additional costs. The rejection of payments and the corresponding additional costs are obviously unjustified, as the provider can nonetheless assign the payment to the right consumers: he informs them that their transfer was rejected.

vzbv, in line with the German Data Ethics Commission (DEK) and the academic community, points out that it is not sufficient to simply refer to GDPR when it comes to the 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 taking decisions about consumers, which leaves consumers unprotected. It

---

[12] Generali Vitality GmbH (see FN. 9).

[13] 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.The "Allianz Schaden Express App" in Austria also automatically decides on cases but usually with a human in the loop Frankfurter Allgemeine Zeitung GmbH: Geld in 30 Sekunden?: Der vollautomatische Kfz-Sachverständige (2018), URL: https://www.faz.net/aktuell/finanzen/meine-finanzen/versichern-und-schuetzen/kuenstliche-intelligenz-in-der-kfz-versicherung-eine-revolution-15374987.html [Access: 20.07.2021]. See also: SVRV - Advisory Council for Consumer Affairs: Consumer-friendly scoring. Report of the Advisory Council for Consumer Affairs (2018), URL: http://www.svr-verbraucherfragen.de/en/ [Access: 20.07.2021]; Insurance Journal: Tokio Marine Uses Tractable's Artificial Intelligence Solution for Auto Claims in Japan (2020), URL: https://www.insurancejournal.com/news/international/2020/05/11/568090.htm [Access: 23.07.2021]

[14] Lemonade: How Lemonade's Tech-Powered Claims Work | Lemonade Insurance (2021), URL: https://www.lemonade.com/de/en/claims [Access: 23.07.2021].

[15] „Another subject for discussion is whether bonus programmes are used as a vehicle for indirect risk selectivity if those who are healthy anyway and those who are health-conscious are the main beneficiaries. The survey shows that some health insurance funds deliberately set out to appeal to health-conscious individuals. This may be interpreted as an attempt to recruit and retain the youngest and healthiest possible clientele.", SVRV - Advisory Council for Consumer Affairs (2018) (see FN. 12), p. 86.

[16] The report of the Advisory Council for Consumer Affairs at Germany's Federal Ministry of Justice and Consumer Protection cites an insurer warning: "Behavioural tariffs may result in individual groups of insured persons exploiting them at the expense of people whose illnesses are not lifestyle-related. We therefore take a very critical view of these tariffs." ebd. p. 86

[17] Compare: Frankenfield, Jake: What Is a Robo-Advisor? in: Investopedia (2021), URL: https://www.investopedia.com/terms/r/roboadvisor-roboadviser.asp [Access: 20.07.2021].

[18] Verbraucherzentrale Bundesverband: Beschwerden zu digitalen Bezahldiensten nehmen zu (2021), URL: https://www.vzbv.de/pressemitteilungen/beschwerden-zu-digitalen-bezahldiensten-nehmen-zu [Access: 04.08.2021].

also becomes increasingly difficult to clearly distinguish between 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. The AIA is a legislation well suited to define minimum legal requirements for profiling and scoring.

AI systems intended to be used in the following areas must be included in the list of high-risk applications in Annex III, because they can a) cause serious economic/financial harm to consumers or 2) severely violate consumer rights:

- AI systems intended to be used in the area of **insurances** including but not limited to AI systems employed to determine individual behaviour-based insurance premiums and rates, risk categorisation of consumers and handling of insurance claims.
- Consumer-facing AI-based applications intended to be used in the area of automated **financial investment or portfolio management**.
- AI-based applications intended to be used in the area of **payment** and **debt collection** services.
- AI systems intended to be used for **scoring and profiling** of consumers when the scores and profiles are used to determine consumers' access to services or markets. The AIA must regulate these AI systems for **scoring and profiling** of consumers as such – and not just decisions based on it, as defined in GDPR.
- AI systems intended to be used to determine access to the **housing market** including but not limited to (pre-)selection of and assigning score values to potential tenants and buyers (if they are consumers, not professionals or legal persons).[19]
- AI systems intended to be used to determine access to **social security services**, including reimbursements.

2.2 Future proofing AIA: Considering consumers rights and economic risks when determining the updating modalities of the list of high-risk applications

Art 7 par. 1. (a) determines that updating the list of high-risk AI systems in Annex III is limited to adding new AI applications within the existing eight areas in Annex III (paragraphs 1-8). Consequently, when updating the list in the future, the European Commission cannot complement the list with high-risk AI applications from other areas of applications. This focus on the present state of AI development endangers the application of this legislation in the future: In the years to come, potentially harmful AI applications could emerge outside the areas listed in Annex III. These would then not fall under the scope of the draft AIA.

To make the AIA future proof, legislators must ensure that they can add new high-risk AI systems as well as areas of application in Annex III outside those eight areas listed in Annex III. Therefore Art. 7 par. 1. (a) must be deleted.

Art. 7 par. 1. (b) holds that when updating the list of high-risk AI applications in Annex III legislators only take AI systems into account that "pose a risk of harm to the health

[19] Discrimination risks in the housing market are various. One example is AI-driven advertisement, potentially discriminatory along racial or religious lines: U.S. Department of Housing and Urban Development: HUD Charges Facebook with Housing Discrimination Over Company's Targeted Advertising Practices (2019), URL: https://www.hud.gov/press/press_releases_media_advisories/HUD_No_19_035 [Access: 29.03.2019].

and safety, or a risk of adverse impact on fundamental rights". There is no reason for not considering violations of consumer rights and economic and financial harms when updating the list in Annex III. These risks can also have significant and long-term negative impact on individuals, groups and their social and economic welfare.

AI-based (personality) profiling of consumers can lead to the exploitation of individual vulnerabilities and the unjustified treatment of persons and groups, leading to economic welfare losses. Examples for AI systems likely leading to unjustified discriminations include AI systems developed to determine consumers' access to markets (e.g. the denial to use platforms or services like AirBnB[20]). Another area of concern are AI systems leading to higher prices in form of unjustified high insurance premiums[21] or rejection of or insurance claims on the basis of faulty AI-driven "lie detectors"[22].

> Art. 7 par. 1. (b) must be complemented so that delegated acts, **updating the list of high-risk AI systems** in Annex III, also take into account the risk of **violation of consumer rights** and **social and economic harms** for individual persons and (social) groups. This will make the AIA future proof, as future AI applications might emerge in new areas of life and significantly affect social and economic welfare of individuals and groups.

3. THE SCOPE IS TOO NARROW PART II: LIST OF PROHIBITED AI

The proposed prohibited practices in Art 5 par. 1 (a), (b), (c) and (d) are too narrow in scope. They leave consumers unprotected from potential harm in various areas. Potential harms to consumers include deception, manipulation und subversion of consumer decisions and their autonomy, leading to economic harm and discrimination as well as potentially substantial violations of privacy.

3.1 Art. 5 par. 1 (a) Prohibition of Dark Patterns

Art. 5 par. 1 (b) aims at prohibiting AI systems exploiting so-called dark patterns[23] to the detriment of consumers. It is well justified and urgent that the AIA addresses the harms caused by dark patterns. 'Dark pattern' refers to unfairly subverting or impairing user autonomy, decision-making, or choice via the structure, function or manner of operation a user interface or a part thereof. Dark pattern-tactics are well-documented and widely used to manipulate users causing financial harm. For example, when they are used to

---

[20] See for example Business Insider: Airbnb has patented software that digs through social media to root out people who display 'narcissism or psychopathy' (2020), URL: https://www.businessinsider.de/international/airbnb-software-predicts-if-guests-are-psychopaths-patent-2020-1/?r=US&IR=T [Access: 12.05.2020].

[21] Compare: McKinsey & Company: Insurance 2030 - The impact of AI on the future of insurance (2021), URL: https://www.mckinsey.com/industries/financial-services/our-insights/insurance-2030-the-impact-of-ai-on-the-future-of-insurance [Access: 20.07.2021]; SVRV - Advisory Council for Consumer Affairs (2018) (see FN. 12).

[22] Quach, Katyanna: Insurance startup backtracks on running videos of claimants through AI lie detector (2021), URL: https://www.theregister.com/2021/05/26/ai_insurance_lemonade/ [Access: 23.07.2021] Bittle, Jake: Lie detectors have always been suspect. AI has made the problem worse., URL: https://www.technologyreview.com/2020/03/13/905323/ai-lie-detectors-polygraph-silent-talker-iborderctrl-converus-neuroid/ [Access: 21.07.2021]

[23] Martini, Mario u. a.: Dark Patterns 01 (2021), in: ZfDR - Zeitschrift für Digitalisierung und Recht, H. 1, URL: https://rsw.beck.de/docs/librariesprovider132/default-document-library/zfdr_heft_2021-01.pdf [Access: 04.05.2021].

keep them from unsubscribing from services[24] or to push gamers during a video game to in-game purchases.[25] About 70% of gamers spent money on in-game purchases. A survey found that 20% of them "spent money without realising that the purchased items would not give them any [..] advantage."[26] AI systems can be exploited to fine tune and personalise manipulative to incentives for continuous in-game spending. Given that in Germany alone gamers spent almost €1.8 billion on in-game purchases in pre-pandemic 2018, the economic welfare losses to European consumers caused by dark patterns via in-game purchases can be expected to be significant.[27]

Art. 5 par. 1 (a) might be well-intended, unfortunately, it will not address many of the most harmful dark patterns as its scope is far too narrow. It prohibits AI systems that "deploy subliminal techniques beyond a person's consciousness" only if they *intentionally* ("in order to") "distort a person's behaviour in a manner that causes or is likely to cause" harm. It is near to impossible to prove that providers of AI systems intended harm, as they will hardly admit to this[28] (e.g. when software for stalking is disguised and marketed as child tracking software[29]). Also "In real life, harm can accumulate without a single event tripping a threshold of seriousness, leaving it difficult to prove. These 'cumulative' harms are reinforced over time by their impact on individuals' environments […]"[30].

Consequently, Art. 5 par. 1 (a) will probably, in practice, not be applicable and enforceable at all.

Art. 5 par. 1 (a) is limited to dark patterns that cause "physical or psychological" harm. But AI systems, by exploiting dark patterns, can cause other significant harms as well, e.g. financial loss, violation of privacy and violation of consumers rights. Consumers should be protected from these kinds of harm as well.

> The AIA should in general **prohibit** that AI systems **exploit dark patterns** by presenting end user choices in a non-neutral manner, or by otherwise subverting or impairing user autonomy, decision-making, or choice via the structure, function or manner of operation of a user interface or a part thereof.

---

[24] For example, the recent Norwegian Consumer Counsel recent report shows how Amazon seems to deliberately obstruct consumers who wish to unsubscribe from its Amazon Prime service. "In the process of unsubscribing from Amazon Prime, the company manipulates consumers to continue using the service in what seems like a deliberate attempt to confuse and frustrate customers." See Forbrukerradet: Amazon manipulates customers to stay subscribed, URL: https://www.forbrukerradet.no/news-in-english/amazon-manipulates-customers-to-stay-subscribed/ [Access: 01.02.2021].

[25] E.g. "A […] federal lawsuit asserts that Electronic Arts unlawfully increases its sports games' difficulty in order to induce gamers into paying the video game publisher additional money." See Sportico: Federal Law Suit: This Video Game is too Damn Hard (2020), URL: https://www.sportico.com/law/analysis/2020/ea-sports-its-in-the-game-1234617287/ [Access: 01.02.2021].

[26] Taylor Wessing LLP: In-game purchases (2019), URL: https://www.taylorwessing.com/download/article-ingame-purchases.html [Access: 21.07.2021].

[27] Germany Trade and Invest - Gesellschaft für Außenwirtschaft und Standortmarketing mbH: Gaming Industry (2021), URL: https://www.gtai.de/gtai-en/invest/industries/creative-industries/gaming-65554 [Access: 21.07.2021].

[28] Compare: Veale, Michael; Borgesius, Frederik Zuiderveen: Demystifying the Draft EU Artificial Intelligence Act, SocArXiv 2021, URL: https://osf.io/preprints/socarxiv/38p5f/download?format=pdf., p.

[29] Harkin, Diarmaid; Molnar, Adam; Vowles, Erica: The commodification of mobile phone surveillance: An analysis of the consumer spyware industry 16 (2020), in: Crime, Media, Culture: An International Journal, H. 1, p. 33–60, URL: https://journals.sagepub.com/doi/10.1177/1741659018820562 [Access: 03.08.2021].

[30] Veale, Michael; Borgesius, Frederik Zuiderveen (see FN. 27), p. 4.

> Art. 5 par. 1 (a) should **not require intentionality** as a precondition for the prohibition of dark patterns, as it is near to impossible to prove. The definition of harm should **not** be limited to physical or psychological harm, but also **include socio-economic welfare losses,** violations of **fundamental rights** (e.g. discrimination) and **consumer rights** (e.g. deception).

3.2 Art. 5 par. 1 (b) Prohibition of exploiting weaknesses to influence behaviour

Just like in Art. 5 par. 1 (a) the European Commission might have good intentions with this article, but due to its narrow scope it will hardly provide any meaningful benefits for consumers in practice.

Art. 5 par. 1 (b) suffers from the same flaw as Art. 5 par.1 (a): It prohibits AI systems exploiting people's weaknesses only if the exploitation happens *intentionally* ("in order to") to materially distort the behaviour which leads to consumer harm. Again, the intention of the provider of the AI system will be near to impossible to prove in practice, even if it existed[31]. Consequently, meaningful application of Art. 5 par. 1 (a) in practice must be doubted.

Also, just like Art. 5 par. 1 (a), Art. 5 par. 1 (b) is limited to exploitations leading to "physical or psychological" harm. However, AI systems, by exploiting vulnerabilities, can cause other significant harms as well, e.g. financial harms, violations of privacy and consumers rights. Consumers should be protect from these kinds of harms as well.

Art 5 par. 1 (b) aims at prohibiting AI systems that exploit a person's or groups weaknesses ("vulnerabilities"), weaknesses of children, the elderly, and the physically or mentally disabled. These groups are certainly particularly vulnerable and worthy of protection. But the problem also exists beyond these groups for other consumers, too.

Every person can be temporarily in a very vulnerable position: emotionally, psychologically or physically. For example, AI systems can exploit vulnerabilities caused by emotional distress, pressure, exhaustion, inattention, tiredness, grief, sorrow, mental agitation, physical pain and injuries or influence of medication or medical treatments.

All consumers should be protected from AI systems exploiting these situational or temporary vulnerabilities. Especially when providers use marketing techniques exploiting consumers' personal - if only temporary - weaknesses in order to overcharge consumers or sell items or services to consumers they would not buy otherwise.[32]

For example, AI systems can be used to personalise and further 'optimise' current techniques for exploitation vulnerable gamers.[33] These include children, but also, problem-

[31] ibid.

[32] Kietzmann, Jan; Paschen, Jeannette; Treen, Emily: Artificial Intelligence in Advertising: How Marketers Can Leverage Artificial Intelligence Along the Consumer Journey 58 (2018), in: Journal of Advertising Research, H. 3, p. 263–267, URL: http://www.journalofadvertisingresearch.com/content/jadvertres/58/3/263.full.pdf [Access: 08.07.2021].

[33] For example in-game purchases. See: Daniel L. King u. a.: Unfair play? Video games as exploitative monetized services: An examination of game patents from a consumer protection perspective 101 (2019), in: Computers in Human Behavior, p. 131–143, URL: https://www.sciencedirect.com/science/article/pii/S0747563219302602 [Access: 21.07.2021].

atic adult gamers with a propensity for game addiction into in-game purchases: “Systems also pair in-game purchase offers with known triggers for an individual player or known triggers for similar players.”[34]

True, also a human salesperson can try to spot customers’ weaknesses and exploit them. But, in contrast to the human, AI-driven marketing systems can draw on large amounts of granular data from various areas in real time to influence consumer decision making: this includes data on individual consumers as well as behavioural patterns of others consumers. Providers can scale up AI systems and systematically exploit individuals’ weaknesses and manipulate many consumers individually[35]. Marketers already use AI systems for real time analysis of consumer behaviour to influence individual decision making along the entire value chain. This ranges from targeted advertisement, to individually curated content and personal rebates influencing consumers’ evaluation and purchase decision[36]. The trend for individualisation of AI-driven marketing will increase even further the existing imbalance of power and knowledge between consumers and providers.

Examples for this AI-driven “hyper-personalisation” of marketing are already in use today are various: E-commerce, websites use AI-driven tools, like Prudsys[37], to analyse user browsing behaviour in real time in order to offer them personalised prices in the form of personalised discounts. In order to push consumers to make purchases, e-Spirit provides AI-driven E-commerce tools to personalise multiple sales channels’ layout, menu bars, displayed ads, pop-ups, text and CTAs (so-called “Calls to action”, meaning designs or phrases intended to prompt an immediate sale[38]). Personalisation can be based on each consumer’s online behaviour or profile. The Canadian car insurer Kanetix used integrate.AI[39] systems to detect and target undecided consumers and increased convergence rate by 13%[40].

Firms can use these tools to exploit consumers’ vulnerabilities like illnesses, exhaustion or other personal difficulties people are struggling with. This can lead to welfare losses when marketers use this technology to overcharge consumers or manipulate them to make purchases they would not do otherwise.

> The AIA should in general **prohibit** AI systems from **exploiting** weaknesses and **vulnerabilities** of **consumers**. This protection should not be limited to young, old and persons with disabilities, but include all consumers, even if their weaknesses or vulnerabilities are temporary or “situational”.

---

[34] Markle, Tracy; Kennedy, Brett: In-Game Purchases: How Video Games Turn Players into Payers. in: Digital Media Treatment (2021), URL: https://digitalmediatreatment.com/in-game-purchases/ [Access: 21.07.2021].

[35] Kietzmann, Jan; Paschen, Jeannette; Treen, Emily (see FN. 31).

[36] Davenport, Thomas u. a.: How Artificial Intelligence Will Change the Future of Marketing 48 (2020), in: Journal of the Academy of Marketing Science, H. 1, p. 24–42, URL: https://link.springer.com/content/pdf/10.1007/s11747-019-00696-0.pdf [Access: 08.07.2021]; Kietzmann, Jan; Paschen, Jeannette; Treen, Emily (see FN. 31); Prudsys: Price Optimization: Intelligent and Personalized Couponing (2021), URL: https://prudsys.de/en/case/price-optimization_promotion-pricing/ [Access: 22.07.2021].

[37] Prudsys (see FN. 35).

[38] Call to action (marketing) - Wikipedia. in: Wikipedia (2021), URL: https://en.wikipedia.org/w/index.php?oldid=1033452251 [Access: 23.07.2021].

[39] Integrate.ai (2021), URL: https://integrate.ai/ [Access: 26.07.2021].

[40] Adriano, Lyle: Kanetix leverages AI technology to optimize consumer experience (2018), URL: https://www.insurancebusinessmag.com/ca/news/digital-age/kanetix-leverages-ai-technology-to-optimize-consumer-experience-93703.aspx [Access: 21.07.2021].

> Art. 5 par. 1 (b) should **not require intentionality** as a precondition for the prohibition of dark patterns, as it is near to impossible to prove. The definition of harm should not be limited to physical or psychological harm, but also include **socio-economic welfare losses**, violations of **fundamental rights** (e.g. discrimination) and violation of **consumer rights** (e.g. deception).

3.3 Art. 5 par. 1 (c) Prohibition of general social scoring by private entities

vzbv welcomes that Art. 5 par. 1 (c) prohibits public authorities from employing AI systems for social scoring under certain conditions.

> The **prohibition** of **social scoring** for the evaluation or classification of people's trustworthiness based on their social behaviour or to **predicted personal** or **personality characteristics** under certain conditions should also include **private entities** and not be limited to scoring undertaken by public authorities. The prohibition of social scoring by private entities should also be subject to the two conditions Art. 5 par. 1 (c) I, namely that the underlying data stems from unrelated social contexts.

General social scoring can have a large negative impact when used by private entities. It can lead to unjustified exclusion of consumers from entire markets or services (or service levels), discrimination and economic, financial, and social harm to consumers or entire groups of consumers.

For example, AI-driven prediction of personality traits from patterns of behaviour collected from smartphone usage[41] or analysis consumers voices (already used in HR contexts), can allegedly reveal personality traits. AI developers provide systems for personality analysis for marketing purposes in E-commerce[42]. A large share of firms say they are eager to employ AI for individually targeting consumers[43].

There are well-funded doubts about the reliability of some currently marketed AI-based systems for personality analysis[44]. Nonetheless, the AIA must regulate these systems,

---

[41] Stachl, Clemens u. a.: Predicting personality from patterns of behavior collected with smartphones 117 (2020), in: Proceedings of the National Academy of Sciences, H. 30, p. 17680–17687, URL: https://www.pnas.org/content/117/30/17680.

[42] E.g. A field of application includes psychological analysis based voice samples, used for HR recruiting context. However, this technique can be applied in consumer contexts as well. Compare Precire: Precire für Marketing und Sales - Zielgerichtete Kundenkommunikation (2021), URL: https://precire.com/marketing-und-sales/ [Access: 21.07.2021]see also Krahmer, Carolin. Recrutainment Blog: Persönlichkeitsprofil aus der Analyse von Sprache: Einfach nur creepy oder die Technologie von morgen? Interview mit Mario Reis von Precire und Britta Nollmann von RANDSTAD (2016), URL: https://blog.recrutainment.de/2016/05/11/persoenlichkeitsprofil-aus-der-analyse-von-sprache-einfach-nur-creepy-oder-die-technologie-von-morgen-interview-mit-mario-reis-von-psyware-und-britta-nollmann-von-randstad/ [Access: 21.07.2021].

[43] A salesforce study reveals the AI's potential for marketers to scale up personalized marketing tools in the area of personalization improvement of lead generation, customer acquisition and upselling. Compare Salesforce: State of Marketing Report - Fifth Edition (2018), URL: https://www.salesforce.com/news/stories/salesforce-releases-fifth-edition-of-state-of-marketing-report-marketers-prioritize-ai-powered-personalization-and-emphasize-customer-trust/ [Access: 19.07.2021].

[44] Compare Thiel, Veronika: Sprachanalyse: Wunschdenken oder Wissenschaft? AlgorithmWatch (2019), URL: https://algorithmwatch.org/de/sprachanalyse-hr/ [Access: 21.07.2021].

in order to be future proof and to prevent discrimination and unfair treatment of consumers due to flawed inferences.

Scoring can be useful for some narrowly defined specific purposes and when subject to strict rules, e.g. credit scoring. Even then it must be subject to strict rules, such as transparency for consumers and scrutiny by independent experts (which is often not the case, as the Schufa credit rating example illustrates[45]).

Both conditions specified in Art 5 par. 1 (c) i and ii seem appropriate to capture the instances when social scoring by private entities becomes overly prone to misjudgements, unjustified outcomes and discrimination: When scoring can lead to detrimental or unfavourable treatment of consumers that is unjustified or disproportioned in social contexts that are unrelated to the contexts in which the data was originally generated or collected. This applies for example to a patent for social scoring hold by AirBnB[46]: Its score aims to predict consumers' trustworthiness based on a variety of social media/online data. This is likely to lead to unjustified discriminatory exclusion and misjudgements of consumers in the vacation housing market. Such misjudgements or discriminations can lead to large individual and aggregated financial and social harm for individual consumers or entire social groups.

3.4 Art. 5 par. 1 (d) Prohibition of remote biometric identification in the public space by private entities

The use of biometric identification systems in publicly accessible spaces can cause significant harm to consumers, including severe violations of the right to privacy and of their autonomy. AI systems that could be used for biometric identification in publicly accessible spaces could occur via smart glasses (e.g. by Facebook[47]) or mobile phone augmented reality applications that can be used to recognise objects in public spaces (e.g. Google Lens[48] or Google maps[49]). These could theoretically also be used for instance to identify passengers in public transport. Other examples are AI systems embedded in cameras in shopping centres.

Whether biometric identification happens in real-time or retrospective often makes no difference with respect to the potential harm e.g. privacy violations or data breaches.[50]

> The AIA should **prohibit** the use of **biometric identification** systems in publicly accessible spaces by **private entities** (not only public authorities). The ban should include 'real-time' as well as retrospective biometric identification.

[45] Compare: AlgorithmWatch: Blackbox Schufa: Auswertung von OpenSCHUFA veröffentlicht -. in: AlgorithmWatch (2018), URL: https://algorithmwatch.org/de/blackbox-schufa-auswertung-von-openschufa-veroeffentlicht/ [Access: 21.07.2021].

[46] Booker beware: Airbnb can scan your online life to see if you're a suitable guest (2020), URL: https://www.standard.co.uk/tech/airbnb-software-scan-online-life-suitable-guest-a4325551.html; Business Insider (see FN. 19).

[47] Compare wearable.com: The best smartglasses and AR specs 2021: Snap, Amazon and more (2021), URL: https://www.wareable.com/ar/the-best-smartglasses-google-glass-and-the-rest [Access: 20.07.2021].

[48] Google: Google Lens – search what you see (2021), URL: https://lens.google/intl/en-GB/ [Access: 20.07.2021].

[49] Dass.: Augmented Reality (2021), URL: https://arvr.google.com/ar/ [Access: 20.07.2021].

[50] For Example Clearview "amassed one of the largest-known repositories of pictures of people's faces — a database of more than 3 billion images scraped without permission from places such as Facebook, Instagram, and LinkedIn." BuzzFeed: Surveillance Nation (2021), URL: https://www.buzzfeednews.com/article/ryanmac/clearview-ai-local-police-facial-recognition [Access: 21.07.2021]. Similar biometric information could be collected from cctv footage or images captured by smart mobile devices in public places.

The use of biometric identification systems by private entities falls under the GDPR, but it should nevertheless be banned outright by the AIA, as the risks to fundamental rights are significant.

> vzbv joins the EDPB's[51] demand that a ban must not be limited to biometric *identification* but that there should be a "general **ban** on any use of AI for **automated recognition of human features** in publicly accessible spaces, such as recognition of faces, gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioural signals, in any context."

In addition to biometric identification and analysis, the collection and analysis of so-called "metabolites"[52] provides particular risks for consumers. It involves the analysis of particles that people leave behind/give off, for example sweat, dust, breath, etc. The analysis of metabolites can allow to draw conclusions about individual behaviour, consumption and habits, and is thus highly sensitive. Therefore, the collection and analysis should be covered by the scope of the AIA regulation.

> The AIA should **ban** the use of AI-based **analysis of "metabolites"** by private entities in consumer facing markets, unless it is to the clear and proven benefit of the consumer (for example health applications in clinical contexts).

3.5 Prohibition of emotion recognition system by private entities

The proposed AIA considers AI-based emotion detection tools a high risk only in the context of law enforcement (this includes polygraphs and similar tools or tools to detect the emotional state of a natural person (Annex III, 6. (b)).

Except for the labelling obligation (Art. 52) which will probably not protect consumers effectively in practice, the draft AIA neglects the issue of emotion detection and analysis in consumer markets.

The automated recognition of human features and expressions (e.g. of faces, mimic, gait, voice) keystrokes and other biometric or behavioural signals by private entities leaves all consumers vulnerable to exploitation, deception and manipulation. Emotion recognition systems can severely harm consumers in commercial contexts for several reasons: Biometric analysis by companies widens the existing asymmetry in information and power between consumers and companies. It greatly enhances the company's position and power in negotiations with consumers.

Companies can exploit biometric emotion recognition to undermine or subvert consumer autonomy and decision-making. This could be done for example by targeting

---

[51] „EDPS call for a general ban on any use of AI for automated recognition of human features in publicly accessible spaces, such as recognition of faces, gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioural signals, in any context.": EDPB & EDPS: EDPB-EDPS Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) (2021), URL: https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-52021-proposal_en.

[52] The Economist: Metabolites and you - People leave molecular wakes that may give away their secrets, in: The Economist, 15.02.2020.

them with personalised offers[53], possibly exploiting current emotional states of consumers (like pain, sorrow, stress, fatigue) to increase the chance of overcharging for goods and services[54]. For example, ShelfPoint[55] introduced real time emotion recognition of facial expressions of customers in retail stores. It assess the shoppers' emotional state as well as demographics like age, gender and ethnic background and ulitmately aims at personalisation of customer engagement via shelf displays increasing convergence[56].

Customer service or sales employees in the markets for high-value goods (e.g. cars) could exploit similar biometric emotion analysis to manipulate consumers and subvert their choices[57]. Insurance Start up lemonade already developed an AI-driven "lie detector". Consumers had to make their claim in form of video. Lemonade's AI would use facial analysis of "'non-verbal cues" to indicate fraudulent consumers claims. In response critique, Lemonade claims it now uses its facial-recognition algorithms to prevent the same person from making multiple claims[58].

In principle, the use of such "lie detectors" could be employed in other areas as well: Imagine landlords interviewing potential tenants using of AI-driven remote "lie detectors"[59] to improve their negation position or put pressure on consumers.

> Art. 5 must be complemented with a provision that bans the use of **AI-based emotion recognition systems** and the analysis of consumers' emotions by **private entities**, except for clearly defined purposes (such as for medical or research purposes in the public interest) in strict compliance with applicable data protection law and subject to appropriate safeguards.

> **Biometric emotion recognition** of consumers should only be **allowed** if it is to the clear and proven **benefit of the consumer**. These include, for example, AI systems in healthcare or medical contexts, such as systems to monitor patients or elderly people in single households. Other useful applications include systems designed to prevent physical harm from consumers, like AI systems that provide clear benefits in related to health or security, like fraud detection.

4. MORE TRANSPARENCY FOR CONSUMERS

The proposal's mandatory transparency requirements towards consumers merely include the labelling of some AI applications (Art 52). Other than this, the draft AIA provides no transparency for consumers, except a CE marking.

---

[53] For example: Kanetix uses AI to categorise and target customers with incentives to buy insurance. Adriano, Lyle (see FN. 39).

[54] See for example O'Shea, Dan: How retailers can tell stories by reading emotions (2018), URL: https://www.retaildive.com/news/how-retailers-can-tell-stories-by-reading-emotions/542298/ [Access: 22.07.2021]; Another Use case for in-store AI systems: Einzelhandelslabor Südwestfalen: Ein Roboter als Kundenberater?, URL: https://www.einzelhandelslabor.de/praxisbeispiele/ein-roboter-als-kundenberater/ [Access: 21.07.2021].

[55] McManus, Ashley: Partner Spotlight: shelfPoint™ Retail Solution Adapts to Shopper Emotions. in: Affectiva (2017), URL: https://blog.affectiva.com/partner-spotlight-shelfpoint-retail-solution-adapts-to-shopper-emotions [Access: 22.07.2021].

[56] Levine, Barry: Cloverleaf's new grocery shelf displays watch shoppers, track their emotions (2017), URL: https://martech.org/cloverleafs-new-grocery-shelf-displays-know-whether-youre-happy/ [Access: 22.07.2021].

[57] Davenport, Thomas, et al. (see FN. 35).

[58] See: Quach, Katyanna (see FN. 21).

[59] Bittle, Jake (see FN. 21).

Legislators must ensure that the AIA includes more requirements for transparency towards consumers. Consumers must obtain the information necessary to make informed decisions and exercise their rights when necessary. They 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 explain to consumers how their systems work to ensure traceability (and accountability).

4.1 Art 52 – Labelling obligation

The labelling obligations in Art. 52 are good in principle[60]. vzbv wants to point out the risk, that the proposed labelling obligations in Art. 52 could be circumvented. For example, the labelling of AI systems interacting with persons or the labelling of emotion recognition systems could become ineffective if the labelling is hidden or hardly recognisable. Also, vzbv points out that producers or providers of deep fake video/audios with malicious intends will probably not comply with the AIA's labelling obligation anyway. Therefore, an effective enforcement of Art. 52 is necessary. The challenge consists of striking a delicate balance between effective enforcement and the right to freedom of expression. Legislators must avoid the misconception of the EU copyright reform that incentivises over-blocking of content by platforms.[61]

> vzbv recognises that the proposed labelling obligations in Art. 52 are important. Legislators must **ensure** that the **labelling obligations** will **not** be **circumvented**.

4.2 Individual explanation for consumers

The draft AIA includes a number of transparency obligations for providers of high-risk AI systems. These include transparency vis-à-vis professional users of the system (Art. 13), supervisory authorities (Art 64) and notified bodies within the context of conformity assessments (annex IV).

Regrettably, there is no provision obliging providers or professional users of AI systems to provide meaningful information to consumers beyond a labelling obligation for a limited set of AI applications (Art 52).

> The AIA must contain a provision mandating **providers of high-risk AI systems** to inform consumers and **explain** the **result** of the individual case in a comprehensible, relevant and concrete manner (upon their request). (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).

> The information must be provided in a comprehensible, relevant and concrete manner and include:
> - The **input data** on the basis of which an AI application made/prepared a decision about the individual. The data must be provided in plain language and a commonly used and machine-readable format. That information must also cover

[60] 85% of German consumers say AI systems should be labelled, see: TÜV-Verband: Sicherheit und Künstliche Intelligenz - Erwartungen, Hoffnungen, Emotionen (2020), URL: https://www.tuev-verband.de/IG-NB/vdtuev-startseite/dok_view?oid=777991 [Access: 21.07.2021].

[61] Verbraucherzentrale Bundesverband: Nutzerrechte sind ein Must-Have - Stellungnahme zum Regierungsentwurf für ein Gesetz zur Anpassung des Urheberrechts an die Erfordernisse des digitalen Binnenmarktes (2021), URL: https://www.vzbv.de/publikationen/nutzerrechte-sind-ein-must-have [Access: 21.07.2021].

the sources from which the data has been obtained and where and by whom the data was originally collected.
- Information about the underlying **logic of the model** and the **criteria** against which the AI system optimises.
- **Measures** to ensure the **fairness**/bias, **robustness**, and **accuracy** of outcomes of the final AI System.
- The **purpose** and goal of the use of the AI system.

Such information rights are central for consumers to be able to understand and individually review an AI system's decision. Only then can consumers 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.[62]

The draft AIA obliges providers of AI systems to generate much of this information in the context of the conformity assessments anyway (e.g. data on robustness and fairness of the model etc.). Therefore, providers will incur no significant extra cost were they obliged to provide this information to consumers, too. This provision is explicitly not aiming at the disclosure of trade secrets but at addressing the legitimate right to information of consumers.

Individual explanations of a high-risk AI application's decision on consumers also provide large benefits for providers of AI systems:

Transparency creates trust among consumers in AI-based systems. This increases consumers' acceptance of the use of these systems in larger areas of life.[63]

Transparency also enables consumers to check the accuracy of AI-based decisions when they can have a significant impact on them personally, e.g. by checking whether a decision about them is based on correct and up-to-date data. This feedback in turn can help improving the accuracy of AI systems' outcomes.

Transparency enables consumers and citizens more widely to exercise their rights and challenge an AI-based decision on a well-founded basis, for example, to defend themselves against discrimination or erroneous decisions. This benefits consumers directly. And consumers taking action against providers of faulty, discriminatory or illegal AI systems provide important benefits to providers of high-quality systems: This exposes the 'black sheep' in the market and redirects demand towards providers who seriously invest in the quality and accuracy of their systems.

4.3 Information for the general public

The implementation of high-risk AI systems can have broad and deep social and economic implications. For example, AI-based selections of job applicants or the determination of insurance premiums can have a profound social and economic impact on the life of many people. With AI systems making or preparing such vital decisions, social

[62] Compare: dass.: White Paper on Artificial Intelligence - Proposals of vzbv (2020), URL: https://www.vzbv.de/sites/default/files/downloads/2020/06/18/20_06_11_vzbv_ec_whitepaper_ai_comment_eng.pdf [Access: 21.07.2021].

[63] For consumers demand for transparency and independent audits of AI systems compare recent surveys by BEUC (see FN. 8); TÜV-Verband (see FN. 59).

trust in AI can only emerge on the basis of an informed public debate and an assessment of the risks and opportunities of these systems. This will in turn encourage uptake and dissemination of AI technologies. The information on high-risk AI systems published in the planned EU database for stand-alone high-risk AI systems (Art. 60) is too superficial to fulfill this purpose[64].

An informed public debate can also serve as guidance for policy-makers regarding the ethical and social implications when deciding about the rules for AI systems in specific sectors or areas of application.

> Providers of high-risk AI systems must provide the **public** with **information** that is relevant for an informed debate and understanding of an AI system. This must entail the information specified in Art 13 par. 3. (b) (characteristics, capabilities and limitations of performance) and Art 13 par. 3. (d) (human oversight).

> The information must be provided in a comprehensible manner and include:
> - General information about the **training data** and the **input data** (e.g. the categories of data).
> - General information about the logic/**methodology** of the **model**.
> - Measures of **fairness/bias**, for the training data and the input/output data (e.g. with respect to gender, ethnicity and other possible grounds of prohibited discrimination).
> - Information and measures on the **robustness** and **accuracy** of the model.

Providers must generate this information anyway to ensure compliance with Art 13. Therefore, this information can be provided at negligible costs. This provision is not aiming at the disclosure of trade secrets but at addressing the legitimate right to information of the general public.

[64] For an overview of the transparency obligations in the draft AIA see Veale, Michael; Borgesius, Frederik Zuiderveen (see FN. 27). P.12

IV. PROPOSALS FOR ENSURING EFFECTIVE INDEPENDENT ASSESSMENTS OF AI SYSTEMS

1. ENSURING CONSUMER TRUST WITH INDEPENDENT ASSESSMENTS

Consumers must be able to trust that all AI systems, especially the high-risk AI systems, comply with all EU legislation when entering EU markets, or when providers change existing systems significantly or employ them in other contexts. Unfortunately, this is currently not the case. Consumers mistrust AI systems, as a recent BEUC[65] consumer survey on perceptions on AI in eight EU Member States shows: Although consumers are generally in favour of AI development, they have serious concerns in relation to AI systems. While consumers see benefits of AI, they have **low trust in AI** and its **added value**. This is displayed in concerns ranging from the lack of transparency, unintended consequences or the abuse of personal data. A **majority** of consumers strongly agree that **companies use AI to manipulate consumer** decisions (e.g. 64 % in Belgium, Italy, Portugal and Spain[66] and even 71 % in Germany[67])

Most consumers think that **current rules** are **not adequate** to effectively **regulate AI**-based activities (50% in Sweden and 55% in Portugal). Around 56% of all EU consumers have low trust in authorities to exert effective control over AI.[68]

> To foster consumers' trust in AI, legislators must **ensure** that **independent** experts can **audit all high-risk AI systems** with respect to compliance with (all) EU legislation and their potential negative impact on consumers.

1.1 Conformity assessment

Legislators cannot leave the assessment of the complex impact of high-risk AI systems to the AI providers' self-assessment. The draft AIA foresees harmonised standards as the basis for the conformity self-assessment. But these standards cannot capture the fine facets of the complex social and economic impact that design and data choices of AI systems have. For example, the social and economic implications and effects of the underlying data, how the data has been collected, interpreted or manipulated (e.g. aggregated, "cleaned" or combined with other data) are highly complex. It requires interpretation and scrutiny from different data and social sciences' perspectives to reveal hidden discrimination of some consumer groups. Legislators must recognise that subtle and indirect forms of discrimination and potential harm to consumers cannot be mitigated by standards in the form a self-assessment.

In the case of most high-risk AI systems, the draft AIA allows that the providers carry out the conformity assessment in the form a self-assessment.[69] Providers must ensure

---

[65] BEUC, 2020, 'Artificial Intelligence: what consumers say: Findings and policy recommendations of a multi-country survey on AI,' https://www.beuc.eu/publications/beuc-x-2020-078_artificial_intelligence_what_consumers_say_report.pdf [download 08.09.2020]

[66] BEUC (see FN. 8).

[67] TÜV-Verband (see FN. 59).

[68] BEUC (see FN. 8).

[69] The exceptions are AI systems for the 'real-time' and 'post' remote biometric identification of natural persons. Here the conformity assessment involves a notified body (see conformity assessment procedure in Annex VII) and the high-

compliance with standards or by “common specifications” and verify these themselves. Consequently, consumers shall trust that the respective AI system complies with the requirements for high-risk AI systems laid out in Title III, Chapter 2.

vzbv points out that self-assessments does not create consumers’ trust, if they cannot be verified by independent auditors. Certification marks relying on self-assessment, like CE markings, are susceptible to misuse and fraud. This is illustrated by cases where producers have affixed CE markings to products that do not fulfil the legal requirements[70], some of which even endangered the health and life of consumers[71]. Therefore it is well justified, that 85% of German consumers say that AI systems should only be brought to the market if their safety has been assessed by independent auditors. Only 17% say that self-assessments by the providers are sufficient.[72] Therefore, the AIA must not leave the conformity assessment of a high-risk system to the providers of high-risk systems.

> Only independent checks and audits of high-risk AI systems can create consumer’s trust and foster the acceptance of AI in general. **All high-risk** applications must be subject to **independently verified conformity assessments** as laid out in Annex VII when
>
> a) the AI system is brought to **market for the first time** and
>
> b) in case there are well funded **indications** that the AI system is **not in conformity** with the requirements in Title III, Chapter 2 any more (e.g. when the system has been changed, or is employed in another context).

> Relying on “common specifications” (Art 41) is not sufficient. Established standards should be the basis for a conformity assessment.[73]

1.2 Title VIII, Chapter 3 – Enforcement must be complemented with independent assessments

vzbv welcomes that the draft AIA ensures that authorities may have access to data, documentation, etc. for monitoring purposes. Art. 64 also allows other public institutions to act as surveillance authorities: “National public authorities or bodies which supervise or enforce the respect of obligations under Union law protecting fundamental rights [...]”. This could for example include Germany’s Federal Anti-Discrimination Agency.

---

risk-AI systems listed in Annex II (e.g. for product safety, medical products etc.), which must adhere to the procedures and external controls laid out in the respective legislation.

[70] See: European Parliament: Answer Given by Mr Verheugen on Behalf of the Commission to Question No P-5938/07 (2008), URL: https://www.europarl.europa.eu/doceo/document/P-6-2007-5938-ASW_EN.html?redirect [Access: 21.07.2021] and TÜV Rheinland: Why Manufacturers Lie About CE Marking (2016), URL: https://insights.tuv.com/blog/why-manufacturers-lie-about-ce-marking [Access: 21.07.2021].

[71] A detailed article on hazards found due to poor-quality AC adapters: "The good news for the consumer is that there appears to be a cheap charger for any make or model of mobile phone, toy or hand-held games consoles that you might require – the bad news is that it could kill you!" Buckinghamshire Trading Standards: “What’s in your socket?” (2008), URL: webarchive.nationalarchives.gov.uk/20140713175508/http%3A/www.buckscc.gov.uk/media/137366/60600_Booklet_proof.pdf [Access: 21.07.2021].

[72] TÜV-Verband (see FN. 59).

[73] DIN - Deutsches Institut für Normung: Standards als zentraler Baustein der europäischen KI-Regulierung (2021), URL: https://www.din.de/de/din-und-seine-partner/presse/mitteilungen/standards-als-zentraler-baustein-der-europaeischen-ki-regulierung-800318 [Access: 21.07.2021].

Civil society organisations have many competences when it comes to the identification of potential risks of AI systems in their specific area of expertise. It would increase trust among consumers if civil society organisations could request and initiate independent assessments by notified bodies on high-risk AI systems if there are reasonable indications that they do not comply with EU legislation or violate consumer rights. This right to request such assessments should be granted to civil society organisations who have the required expertise. They should include human rights- labour-, consumer- and environmental organisations.

Such audits would provide an independent assessment of a high-risk AI system. They could include an assessment of the legality of the AI-System (e.g. with respect to discrimination or consumer rights), their social, economic and environmental risks and benefits as well as their impact on the psychological and physical health of individual persons, social groups and society.

Such independent assessments can be the basis for an informed public discussion on the risks and benefits of a high-risk AI System and serve as a watchdog. Independent auditors scrutinising high-risk AI systems on the initiative of civil society organisations will increase people's trust in the respective AI systems.[74]

> Limiting **market surveillance** to public **authorities** and institutions as Art. 64 par. 3 is **not sufficient**. **Civil society organisations** must have the right to **request audits** of high-risk AI systems by notified bodies when there are reasonable indications that the high-risk AI system violates European or Member States' legislation, has a significant negative impact on the social, economic, physical or psychological well-being or the security of persons or social groups, or poses significant environmental risks.

> Legislators must complement the draft AIA with **due process obligations** for providers of high-risk AI systems so that notified bodies, on the request of civil society organisations, can conduct independent audits. This must include obligations to give the auditors access to all data, documentation and records (as laid out in Art. 64 par. 1) needed by the auditors to assess the high-risk AI systems' risks to the social, economic, physical or psychological wellbeing and security of persons or groups as well as its potential environmental impact.

> The notified body must publish the findings of the **audit** in a **report**. Policy-makers should establish safeguards, ensuring that confidential information is protected (for example via confidentiality agreements). However, policy-makers must ensure that providers of AI applications do not advance the protection of trade secrets for withholding information from scrutiny by the auditors or exclude it from the published report.

[74] Consumers' demand for independent scrutiny of AI systems see: BEUC (see FN. 8); TÜV-Verband (see FN. 59).

V. ENSURING PRIVATE ENFORCEMENT

Private enforcement of EU legislation complements the enforcement efforts by competent authorities. Consumers greatly benefit when consumer organisations enforce their rights in courts complementary to enforcement by competent authorities:[75] Consumer organisations are well aware of the detriments consumers are facing in various markets. Consumer organisations like vzbv can take proactive action. In doing so, they prevent consumer harm from occurring in the first place and avoid disputes as far as possible. The European Commission points to the large number of injunction procedures in Germany and Austria "which both traditionally rely on the private enforcement of consumer law initiated by the consumer and business organisations".[76] As a result, one can expect a reduction in the number of infringements by companies, leading to a reduction in related consumer detriment.[77] For example, more than half of vzbv's legal actions are successfully settled out of court with companies issuing cease-and-desist declarations.[78] vzbv's successful procedures against Volkswagen resulted in a €830 million settlement for German consumers.[79] Also, vzbv proceedings against Facebook[80] illustrate the benefits of consumer organisations enforcing consumer law. In addition, enforcement of consumer rights by consumer organisations relieves the enforcement burden on competent authorities. It frees up authorities' resources allowing them to concentrate scarce resources on strategically important cases.

> To ensure that consumer organisations can enforce the AIA provisions, legislators must add the AIA to Annex I of the European Directive on representative actions for the protection of the collective interests of consumers ((EU) 2020/1828).[81]

---

[75] Verbraucherzentrale Bundesverband: Mehr Sammelklage wagen - Kurzpapier des vzbv (2021), URL: https://www.vzbv.de/pressemitteilungen/mehr-sammelklage-wagen [Access: 21.07.2021]

[76] European Commission: Report of the Fitness Check SWD(2017)209, URL: https://ec.europa.eu/transparency/documents-register/detail?ref=SWD(2017)209&lang=en [Access: 14.07.2021].

[77] Ebd. p.103

[78] Verbraucherzentrale Bundesverband: Broschüre: Recht durchsetzen, Verbraucher stärken (2015), URL: https://www.vzbv.de/publikationen/broschuere-recht-durchsetzen-verbraucher-staerken [Access: 14.07.2021].

[79] See: Volkswagen AG: European Directive on representative actions for the protection of the collective interests of consumers ((EU) 2020/1828) (2020), URL: https://www.volkswagenag.com/en/news/2020/02/vzbv-and-volkswagen-agree-on-a-fair-settlement-solution.html# [Access: 20.07.2021]; Verbraucherzentrale Bundesverband: vzbv-Klage gegen VW führt zu Deutschlands größtem Massenvergleich (2020), URL: https://www.vzbv.de/urteile/vzbv-klage-gegen-vw-fuehrt-zu-deutschlands-groesstem-massenvergleich [Access: 21.07.2021].

[80] Verbraucherzentrale Bundesverband: Facebook verstößt gegen Datenschutzrecht - Kammergericht Berlin gibt Klag also pair in-game purchase offers with known triggers e des vzbv in vielen Punkten statt (2020), URL: https://www.vzbv.de/urteile/facebook-verstoesst-gegen-datenschutzrecht [Access: 21.07.2021].

[81] European Parliament: Directive (EU) 2020/1828 of the European Parliament and of the Council of 25 November 2020 on representative actions for the protection of the collective interests of consumers and repealing Directive 2009/22/EC (2020), URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32020L1828 [Access: 21.07.2021].

VI. TRADE AGREEMENTS MUST NOT HINDER AN EFFECTIVE TRANSPARENCY AND MONITORING OF AI SYSTEMS

vzbv wants to highlight the importance of consistency of European AI-related policies. It is of particular importance to ensure the compatibility of the AIA with trade commitments to which the EU is binding itself by international law, especially as AI technologies and the understanding of risks is still nascent and will likely be evolving in the years to come. A vzbv study[82] recently found that current EU trade negotiations might significantly restrict the EU's ability to regulate in the field of AI in the future, in particular with regard to independent assessments and audits. The study finds that trade rules could impede on future EU rules on transparency, certification and accountability. Potential rules on the non-disclosure of source code currently under discussion in the World Trade Organisation (WTO) would hinder effective transparency provisions within the AIA.

> Legislators must enact **trade rules** that do **not impede** on future **AIA rules** on transparency, certification and accountability.
>
> Potential rules on the non-disclosure of source code currently under discussion in the World Trade Organisation (WTO) must not hinder an effective transparency, enforcement, monitoring and independent assessments of AI systems under the AIA.

[82] Irion, Kristina: AI Regulation in the European Union and Trade Law: How Can Accountability of AI and a High Level of Consumer Protection Prevail over a Trade Discipline on Source Code? (2021), URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3786567 [Access: 21.07.2021].