**AI ACT – FEDMA’S INPUT FOR THE TRILOGUE DISCUSSIONS**

**INDEX**

1. [**Recital 16 on Prohibited Artificial Intelligence Practices**](#Recital16)
2. [**Article 3(1)(1) on Definitions**](#DEFINITIONS)
3. [**Article 5(1)(a) on Prohibited Artificial Intelligence Practices**](#Prohibited)
4. [**Article 6 on Classification rules for high-risk AI systems**](#classification)
5. [**Article 28(b) new on Obligations of the provider of a foundation model**](#foundation)
6. [**Annex III on High-Risk AI Systems referred to in Article 6(2)**](#annexIII)

**TABLE OF COMPARISON [[1]](#footnote-2)**

|  |  |  |  |
| --- | --- | --- | --- |
| [**COMMISSION**](https://data.consilium.europa.eu/doc/document/ST-8115-2021-INIT/en/pdf)[**(Annexes)**](https://eur-lex.europa.eu/resource.html?uri=cellar:e0649735-a372-11eb-9585-01aa75ed71a1.0001.02/DOC_2&format=PDF) | [**PARLIAMENT**](https://88fe0205-a2ac-4895-ba2e-a260b7a7b33d.usrfiles.com/ugd/88fe02_d1ead08db89d4b259552ec545415b1d4.pdf) | [**COUNCIL**](https://data.consilium.europa.eu/doc/document/ST-14954-2022-INIT/en/pdf) | **Comments** |
| **RECITAL 16** |
| The placing on the market, putting into service or use of certain AI systems intended to distort human behaviour, whereby physical or psychological harms are likely to occur, should be forbidden. Such AI systems deploy subliminal components individuals cannot perceive or exploit vulnerabilities of children and people due to their age, physical or mental incapacities. They do so with the intention to materially distort the behaviour of a person and in a manner that causes or is likely to cause harm to that or another person. The intention may not be presumed if the distortion of human behaviour results from factors external to the AI system which are outside of the control of the provider or the user. Research for legitimate purposes in relation to such AI systems should not be stifled by the prohibition, if such research does not amount to use of the AI system in human machine relations that exposes natural persons to harm and such research is carried out in accordance with recognised ethical standards for scientific research. | The placing on the market, putting into service or use of certain AI systems with the objective to or the effect of materially distorting human behaviour, whereby physical or psychological harms are likely to occur, should be forbidden. This limitation should be understood to include neuro-technologies assisted by AI systems that are used to monitor, use, or influence neural data gathered through brain-computer interfaces insofar as they are materially distorting the behaviour of a natural person in a manner that causes or is likely to cause that person or another person significant harm. Such AI systems deploy subliminal components individuals cannot perceive or exploit vulnerabilities of individuals and specific groups of persons due to their known or predicted personality traits, age, physical or mental incapacities, social or economic situation. They do so with the intention to or the effect of materially distorting the behaviour of a person and in a manner that causes or is likely to cause significant harm to that or another person or groups of persons, including harms that may be accumulated over time. The intention to distort the behaviour may not be presumed if the distortion results from factors external to the AI system which are outside of the control of the provider or the user, such as factors that may not be reasonably foreseen and mitigated by the provider or the deployer of the AI system. In any case, it is not necessary for the provider or the deployer to have the intention to cause the significant harm, as long as such harm results from the manipulative or exploitative AI-enabled practices. The prohibitions for such AI practices is complementary to the provisions contained in Directive 2005/29/EC, according to which unfair commercial practices are prohibited, irrespective of whether they carried out having recourse to AI systems or otherwise. In such setting, lawful commercial practices, for example in the field of advertising, that are in compliance with Union law should not in themselves be regarded as violating prohibition. Research for legitimate purposes in relation to such AI systems should not be stifled by the prohibition, if such research does not amount to use of the AI system in humanmachine relations that exposes natural persons to harm and such research is carried out in accordance with recognised ethical standards for scientific research and on the basis of specific informed consent of the individuals that are exposed to them or, where applicable, of their legal guardian. | The harmonised rules on the placing on the market, putting into service and use of AI systems laid down in this Regulation should apply across sectors and, in line with its New Legislative Framework approach, should be without prejudice to existing Union law, notably on data protection, consumer protection, fundamental rights, employment and product safety, to which this Regulation is complementary. As a consequence all rights and remedies afforded by such Union law to consumers and other persons who may be negatively impacted by AI systems, including as regards the compensation of possible damages pursuant to Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products, remain unaffected and fully applicable. On top of that, this Regulation aims to strengthen the effectiveness of such existing rights and remedies by establishing specific requirements and obligations, including in respect of transparency, technical documentation and record-keeping of AI systems. Furthermore, the obligations placed on various operators involved in the AI value chain under this Regulation should apply without prejudice to national laws, in compliance with Union law, having the effect of limiting the use of certain AI systems where such laws fall outside the scope of this Regulation or pursue other legitimate public interest objectives than those pursued by this Regulation. For example, national labour law and the laws on the protection of minors (i.e. persons below the age of 18) taking into account the United Nations General Comment No 25 (2021) on children’s rights, insofar as they are not specific to AI systems and pursue other legimitate public interest objectives, should not be affected by this Regulation. | **KEY CHANGES**The EP inserted a new reference in Recital 16 to commercial practices that are lawful under the [Unfair Commercial Practices Directive](https://easa-alliance.us2.list-manage.com/track/click?u=cb8fee4e8a9cd3fc5cd878638&id=dfbc6de4ff&e=d6df126be2) – such as advertising lawfully having recourse to AI, which would not be prohibited because of them resorting to AI.**FEDMA’S RECOMMENDATION:**Avoiding a broad definition of AI which could encompass many software applications which could lead to "retroactive" costs to bring systems already in the market into compliance.**WHICH TEXT DO YOU SUPPORT?*** European Commission
* European Parliament
* Council of the EU
 |
| **ARTICLE 3(1)(1)****Definitions** |
| (1) ‘artificial intelligence system’ (AI system) means software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with; | (1) ‘‘artificial intelligence system’ (AI system) means *~~software~~* a machine-based system that is *~~developed~~* designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as *~~content~~*; predictions, recommendations, or decisions, that influence physical or virtual environments *~~they interact with;~~* | (1) ‘artificial intelligence system’ (AI system) means a *~~software~~* system that is *~~developed~~* designed to operate with elements of autonomy and that, based on machine and/or human-provided data and inputs, infers how to achieve a given set of objectives using machine learning and/or logic- and knowledge based approaches, and produces system-generated outputs such as content (generative AI systems), predictions, recommendations or decisions, influencing the environments with which the AI system interacts *~~they interact with~~*; | **KEY CHANGES*** Both the EP and the Council aim to align their definition with the OECD's.
* The EP's definition describes an AI system as a machine-based system that operates with varying levels of autonomy and can generate output that influences physical or virtual environments.
* The Council narrowed down the definition to focus on machine learning and logic- and knowledge-based approaches, excluding traditional software.

**FEDMA’S RECOMMENDATION:**Avoiding a broad definition of AI which could encompass many software applications which could lead to "retroactive" costs to bring systems already in the market into compliance.**WHICH TEXT DO YOU SUPPORT?*** European Commission
* European Parliament
* Council of the EU
 |
| **ARTICLE 5(1)(a)****Prohibited Artificial Intelligence Practices** |
| (a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person’s consciousness in order to materially distort a person’s behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm; | (a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person’s consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person’s or a group of persons’ behaviour by appreciably impairing the person’s ability to make an informed decision, thereby causing the person to take a decision that that person would not have otherwise taken in a manner that causes or is likely to cause that person, another person or group of persons significant *~~physical and psycological~~ harm;*The prohibition of AI system that deploys subliminal techniques referred to in the first sub-paragraph shall not apply to AI systems intended to be used for approved therapeutical purposes on the basis of specific informed consent of the individuals that are exposed to them or, where applicable, of their legal guardian; | (a) the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person’s consciousness with the objective to or the effect of materially distorting a person’s behaviour in a manner that causes or is reasonably likely to cause that person or another person physical or psychological harm; | **KEY CHANGES**Harm:The EP removes the distinction between physical/psychological harm while referring to significant harm à broad meaning. The Council maintains such distinction. Dark Patterns:The EP explicitly refers to the use of dark patterns when referring to manipulative/deceptive techniques**FEDMA’S RECOMMENDATION:**Making sure that the wording of the Article does not leave for interpretation suggesting that legitimate marketing practices, powered by AI, for the promotion of products and services could be considered as subliminal, deceptive or manipulative techniques beyond a person’s consciousness.**WHICH TEXT DO YOU SUPPORT?*** European Commission
* European Parliament
* Council of the EU
 |
| **ARTICLE 6****Classification rules for high-risk AI systems** |
| Irrespective of whether an AI system is placed on the market or put into service independently from the products referred to in points (a) and (b), that AI system shall be considered high-risk where both of the following conditions are fulfilled: (a) the AI system is intended to be used as a safety component of a product, or is itself a product, covered by the Union harmonisation legislation listed in Annex II; (b) the product whose safety component is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation legislation listed in Annex II. 2. In addition to the high-risk AI systems referred to in paragraph 1, AI systems referred to in Annex III shall also be considered high-risk. | Irrespective of whether an AI system is placed on the market or put into service independently from the products referred to in points (a) and (b), that AI system shall be considered high-risk where both of the following conditions are fulfilled: 1. the AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonisation law listed in Annex II;
2. the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment related to risks for health and safety, with a view to the placing on the market or putting into service of that product pursuant to the Union harmonisation law listed in Annex II;

2. In addition to the high-risk AI systems referred to in paragraph 1, AI systems falling under one or more of the critical areas and use cases referred to in Annex III shall be considered high-risk if they pose a significant risk of harm to the health, safety or fundamental rights of natural persons. Where an AI system falls under Annex III point 2, it shall be considered to be high-risk if it poses a significant risk of harm to the environment. The Commission shall, six months prior to the entry into force of this Regulation, after consulting the AI Office and relevant stakeholders, provide guidelines clearly specifying the circumstances where the output of AI systems referred to in Annex III would pose a significant risk of harm to the health, safety or fundamental rights of natural persons or cases in which it would not.2 a. Where providers falling under one or more of the critical areas and use cases referred to in Annex III consider that their AI system does not pose a significant risk as described in paragraph 2, they shall submit a reasoned notification to the national supervisory authority that they are not subject to the requirements of Title III Chapter 2 of this Regulation. Where the AI system is intended to be used in two or more Member States, that notification shall be addressed to the AI Office. Without prejudice to Article 65, the national supervisory authority shall review and reply to the notification, directly or via the AI Office, within three months if they deem the AI system to be misclassified.2 b. Providers that misclassify their AI system as not subject to the requirements PE731.563v02-00 146/657 RR\1279290EN.docx EN of Title III Chapter 2 of this Regulation and place it on the market before the deadline for objection by national supervisory authorities shall be subject to fines pursuant to Article 71.2 c. National supervisory authorities shall submit a yearly report to the AI Office detailing the number of notifications received, the related highrisk areas at stake and the decisions taken concerning received notifications | 1.An AI system that is itself a product covered by the Union harmonisation legislation listed in Annex II shall be considered as high risk if it is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to the above mentioned legislation.2. An AI system intended to be used as a safety component of a product covered by the legislation referred to in paragraph 1 shall be considered as high risk if it is required to undergo a third-party conformity assessment with a view to the placing on the market or putting into service of that product pursuant to above mentioned legislation. This provision shall apply irrespective of whether the AI system is placed on the market or put into service independently from the product. 3. AI systems referred to in Annex III shall be considered high-risk unless the output of the system is purely accessory in respect of the relevant action or decision to be taken and is not therefore likely to lead to a significant risk to the health, safety or fundamental rights. In order to ensure uniform conditions for the implementation of this Regulation, the Commission shall, no later than one year after the entry into force of this Regulation, adopt implementing acts to specify the circumstances where the output of AI systems referred to in Annex III would be purely accessory in respect of the relevant action or decision to be taken. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 74, paragraph 2. | KEY CHANGES:* The Council introduced an "extra criterium" to determine whether an AI system is high risk based on the significance of the AI system's output in the decision-making process, in order to exclude purely accessory outputs.
* The EP proposed a self-assessment system where AI developers determine if their application is high risk using guidance from the EU Commission. If a company believes their system is not high risk, they must inform the relevant authority, which has three months to respond if they disagree with the classification.

**FEDMA’S RECOMMENDATION:**Refraining from categorizing AI systems used in data-driven marketing as high-risk: currently only AI systems used for advertising job vacancies are included in the list by all EU institutions.**WHICH TEXT DO YOU SUPPORT?*** European Commission
* European Parliament
* Council of the EU
 |
| **ARTICLE 28(b) new****Obligations of the provider of a foundation model** |
|  | 1. A provider of a foundation model shall, prior to making it available on the market or putting it into service, ensure that it is compliant with the requirements set out in this Article, regardless of whether it is provided as a standalone model or embedded in an AI system or a product, or provided under free and open source licences, as a service, as well as other distribution channels.2. For the purpose of paragraph 1, the provider of a foundation model shall: (a) demonstrate through appropriate design, testing and analysis that the identification, the reduction and mitigation of reasonably foreseeable risks to health, safety, fundamental rights, the environment and democracy and the rule of law prior and throughout development with appropriate methods such as with the involvement of independent experts, as well as the documentation of remaining non-mitigable risks after development(b) process and incorporate only datasets that are subject to appropriate data governance measures for foundation models, in particular measures to examine the suitability of the data sources and possible biases and appropriate mitigation PE731.563v02 -00 206 /657 RR\1279290EN.docx EN (c) design and develop the foundation model in order to achieve throughout its lifecycle appropriate levels of performance, predictability, interpretability, corrigibility, safety and cybersecurity assessed through appropriate methods such as model evaluation with the involvement of independent experts, documented analysis, and extensive testing during conceptualisation, design, and development; (d) design and develop the foundation model, making use of applicable standards to reduce energy use, resource use and waste, as well as to increase energy efficiency, and the overall efficiency of the system, whithout prejudice to relevant existing Union and national law. This obligation shall not apply before the standards referred to in Article 40 are published. Foundation models shall be designed with capabilities enabling the measurement and logging of the consumption of energy and resources, and, where technically feasible, other environmental impact the deployment and use of the systems may have over their entire lifecycle; (e) draw up extensive technical documentation and intelligible instructions for use, in order to enable the downstream providers to comply with their obligations pursuant to Articles 16 and 28(1); . (f) establish a quality management system to ensure and document compliance with this Article, with the possibility to experiment in fulfilling this requirement, (g) register that foundation model in the EU database referred to in Article 60, in accordance with the instructions outlined in Annex VIII point C. When fulfilling those requirements, the generally acknowledged state of the art shall be taken into account, including as reflected in relevant harmonised standards or common specifications, as well as the latest assessment and measurement methods, reflected in particular in benchmarking guidance and capabilities referred to in Article 58a; 3. Providers of foundation models shall, for a period ending 10 years after their foundation models have been placed on the market or put into service, keep the technical documentation referred to in paragraph 2(e) at the disposal of the national competent authorities 4. Providers of foundation models used in AI systems specifically intended to generate, with varying levels of autonomy, content such as complex text, images, audio, or video (“generative AI”) and providers who specialise a foundation model into a generative AI system, shall in addition a) comply with the transparency obligations outlined in Article 52 (1), b) train, and where applicable, design and develop the foundation model in such a way as to ensure adequate safeguards against the generation of content in breach of Union law in line with the generally-acknowledged state of the art, and without prejudice to fundamental rights, including the freedom of expression,c) without prejudice to Union or national or Union legislation on copyright, document and make publicly available a sufficiently detailed summary of the use of training data protected under copyright law. |  | KEY CHANGES:* Providers of foundation models would be required to ensure the robust protection of fundamental rights, health and safety, environment, democracy, and the rule of law.
* They must assess and mitigate risks, adhere to design, information, and environmental requirements, and register in the EU database.
* Generative foundation models like GPT must meet additional transparency requirements, such as disclosing that the content was generated by AI, designing the model to prevent the generation of illegal content, and publishing summaries of copyrighted data used for training.

FEDMA’S TAKEAWAYSAs foundation models are increasingly being used in data-driven marketing, there is a risk that the cost of compliance with excessive requirements by providers of these models would be passed down to their users (e.g. marketers).**TO WHAT EXTENT MAY THE EP’S TEXT AFFECT FOUNDATION MODELS EMPLOYED IN THE DATA & MARKETING SECTOR?** |
| **ANNEX III****High-Risk AI Systems referred to in Article 6(2)** |
| High-risk AI systems pursuant to Article 6(2) are the AI systems listed in any of the following areas: 1. Biometric identification and categorisation of natural persons: (a) AI systems intended to be used for the ‘real-time’ and ‘post’ remote biometric identification of natural persons; 2. Management and operation of critical infrastructure […]3. Education and vocational training […] 4. Employment, workers management and access to self-employment: (a) AI systems intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests; (b) AI intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behavior of persons in such relationships. 5. Access to and enjoyment of essential private services and public services and benefits […] 6. Law enforcement […] (7. Migration, asylum and border control management […]8. Administration of justice and democratic processes […] | The AI systems specifically refered to in under points 1 to 8a stand for critical use cases and are each considered to be highrisk AI systems pursuant to Article 6(2), provided that they fulfil the criteria set out in that Article:1.Biometric and biometrics-based systems(a) AI systems intended to be used for *~~the ‘real-time’ and ‘post’ remote~~* biometric identification of natural persons, with the exception of those mentioned in Article 5;(a a) AI systems intended to be used to make inferences about personal characteristics of natural persons on the basis of biometric or biometrics-based data, including emotion recognition systems, with the exception of those mentioned in Article 5; Point 1 shall not include AI systems intended to be used for biometric verification whose sole purpose is to confirm that a specific natural person is the person he or she claims to be.1. Management and operation of critical infrastructure […]
2. Education and vocational training […]

4. Employment, workers management and access to self employment: (a) AI systems intended to be used for recruitment or selection of natural persons, notably for placing targeted job advertisements screening or filtering applications, evaluating candidates in the course of interviews or tests;[…]8. Administration of justice and democratic processes: […](a a) AI systems intended to be used for influencing the outcome of an election or referendum or the voting behaviour of natural persons in the exercise of their vote in elections or referenda. This does not include AI systems whose output natural persons are not directly exposed to, such as tools used to organise, optimise and structure political campaigns from an administrative and logistic point of view.(a b) AI systems intended to be used by social media platforms that have been designated as very large online platforms within the meaning of Article 33 of Regulation EU 2022/2065, in their recommender systems to recommend to the recipient of the service user generated content available on the platform. | In each of the areas listed under points 1-8, the AI systems specifically mentioned under each letter are considered to be high-risk AI systems pursuant to Article 6(3): 1. Biometrics: (a) Remote biometric identification systems. 2. Critical infrastructure […]3. Education and vocational training: […]4. Employment, workers management and access to self-employment: (a) AI systems intended to be used for recruitment or selection of natural persons, notably to place targeted job advertisements, to analyse and filter job applications, and to evaluate candidates;(b) AI intended to be used to make decisions on promotion and termination of workrelated contractual relationships, to allocate tasks based on individual behavior or personal traits or characteristics and to monitor and evaluate performance and behavior of persons in such relationships. 5. Access to and enjoyment of essential private services and essential public services and benefits: […]6. Law enforcement: […]7. Migration, asylum and border control management: […]. 8. Administration of justice and democratic processes: […] | **KEY CHANGES*** The Council removed from the list of high risk AI deep fake detection by law enforcement, crime analytics, and verification of travel document authenticity. They added critical digital infrastructure and life and health insurance to the list.
* The EP expanded the list by introducing biometrics, critical infrastructure, recommender systems used by major social media platforms, AI systems that may impact electoral outcomes, AI used in dispute resolutions, and border management.

**FEDMA’S RECOMMENDATION:**Refraining from categorizing AI systems used in data-driven marketing as high-risk: currently only AI systems used for advertising job vacancies are included in the list by all EU institutions.**WHICH TEXT DO YOU SUPPORT?*** European Commission
* European Parliament
* Council of the EU
 |

1. *~~Text:~~* deleted from the EC proposal

Text: Added in the adopted negotiating text

*~~Text~~*: Deleted by both the EP and the Council

Text: Added by both the EP and the Council [↑](#footnote-ref-2)