# Position Paper

## The EU Artificial Intelligence Act – Bitkom Position for the trilogue negotiations
13.7.2023

**Kai Pascal Beerlink**
Policy Officer AI

T +49 30 27576-278
k.beerlink@bitkom.org

**Lukas Klingholz**
Head of Cloud & AI

T +49 30 27576-101
l.klingholz@bitkom.org

**Lea Ossmann-Magiera**
Research Associate AI

T +49 30 27576-181
l.ossmann@bitkom.org

Albrechtstraße 10
10117 Berlin

## Summary

In April 2021, the European Commission (EU COM) presented its proposal for harmonised rules on Artificial Intelligence, the Artificial Intelligence Act ("AIA"). After the Council of the European Union ("Council") and the European Parliament ("EP") have finalised their positions in December 2022 and June 2023, the law-making institutions have entered the interinstitutional negotiations ("trilogue") on the AIA in June 2023.

With the proposed regulation, European co-legislators aim to foster trust in AI and protect fundamental rights while enabling innovation and enhancing Europe's competitiveness at the same time. So far, there is no agreement on how these goals are to be achieved. Since the Council's and the EP's amendments differ strongly from the EU COM's proposal in some crucial points**, the upcoming trilogue negotiations will play a decisive role** in determining what the AIA will look like in concrete terms and whether a fair balance will be struck between protecting against risks and seizing opportunities.

For this reason, Bitkom once again takes the opportunity to comment on the legislative proposals by expressing support and concerns regarding the AIA's most important points. As a cross-sectional technology, AI will continuously extend its impact on all areas of life. It is, hence, of utmost importance to design **a reasonable, future-proof, and proportionate regulatory framework** for the development, distribution, and use of AI systems.

Especially with regard to the **definition of AI**, the **risk classification**, the distribution of **roles and responsibilities** under the AIA, the inclusion of so-called **General Purpose Artificial Intelligence ("GPAI")** systems and the **interplay between the AIA and existing laws**, a reasonable compromise ought to be found. Our main points on these aspects are the following:

- Regulate what makes AI special by concentrating on methods that learn from data[1].
- Create a truly risk-based approach by adding an additional criterion of 'significant risk of harm' to the risk-classification
- Make sure that interaction with sectorial legislation is frictionless by having consistent definitions and by avoiding duplication of requirements.
- Focus on cooperation within the value chain to effectively address risks instead of creating requirements for certain AI-systems that do not have an intended purpose.
- Strengthen the promotion and not only the restriction of AI by creating effective regulatory sandboxes.
- Make the fast development of harmonised standards a strategic top priority.

[1] A definition of an AI system should include certain degrees of autonomy, and non-rule based systems that generate an output. In this context it's important to have as clear definitions as possible.

# Content

# Introduction

With this paper, we would like to take the opportunity once again to address the most important aspects of the proposed AIA and to position ourselves on its central provisions. We appeal to the part

Parties involved in the trilogue negotiations to work towards a proportionate regulation of AI that is capable of fostering trust in the technology. In our view, this goal can only be achieved if, on the one hand, a secure development of AI systems is guaranteed and, on the other hand, the economic use of AI technologies within Europe is enabled and promoted by law.

Hence, when deciding upon an upcoming legal framework for AI, particular attention should be paid to the following aspects:

# Subject matter, scope, and approach

## Subject matter

In art. 1 of their AIA proposal, the EU COM mentions the harmonisation of rules for AI systems, the prohibition of certain AI practices, specific requirements for high-risk systems, transparency rules, and market monitoring and surveillance as the subject matter of the AIA.

In contrast to the EU COM's proposal, the Council's approach added 'measures in support of innovation' as a subject matter of the AIA to art. 1 (e).

The EP has added measures to support innovation, with a particular focus on SMEs and start-ups, including setting up regulatory sandboxes and targeted measures to reduce the regulatory burden on SMEs and start-ups to the scope (art. 1 (ea)).

**Bitkom position**

We welcome these amendments, since it emphasizes the importance of fostering innovation – one goal that specifically with regard to AI has tremendous relevance for the European Union and, thus, should be equally addressed by the legislator. Since the subject matter of a regulation can be used for legal interpretation of indefinite legal terms, this amendment will become relevant once the AIA has entered into force.

## Material scope

The EU COM's proposal excludes AI systems developed or used exclusively for military purposes and AI systems used by public authorities in a third country or international organisations from the AIA's scope.

The Council's approach adds exceptions for AI used solely for scientific research and development or for purely personal non-professional activity (except for art. 52 – transparency obligation).

The EP has added that the AIA shall not preclude Member States or the Union from maintaining or introducing laws, regulations or administrative provisions which are more favourable to workers in terms of protecting their rights in respect of the use of AI systems by employers, or to encourage or allow the application of collective agreements which are more favourable to workers (art. 2 (5c)).

**Bitkom position**

Regulating AI on a European Level supports the single market approach. Bitkom sees the introduction of an opening clause  as proposed by the EP therefore critical, as it would undermine this original thought.

Also, leaving the possibility to impose more restrictive regulations could have the effect of slowing down innovation and development.

## Risk-based approach and risk classification

Bitkom welcomes the EU COM proposal's risk-based approach to AI regulation. With regard to the principle of proportionality, the risk-based approach which has been taken by the EU COM has not been (seriously) challenged during the regulatory process. It is, however, essential to limit the high-risk category to those systems that actually pose a high risk to fundamental rights or health and safety, and not to regulate technologies generally. Since manufacturers of high-risk systems face a wide range of obligations (see below), it would be fatal to wrongly place AI systems in the high-risk category when in fact they do not pose a high risk. Hence, the **risk classification is one cornerstone of AI regulation** which must be addressed seriously in the trilogue negotiations.

### Risk classification under the EU COM's proposal

The risk classification is carried out through an interplay of art. 6 with the Annexes. This mechanism proposed by the EU COM has not been amended by the co-legislators.

According to art. 6, AI systems are considered high risk if they are covered by the legislative framework listed in Annex II and if, according to this framework, the system must undergo a third-party conformity assessment before being placed on the market.  The same provision is proposed for safety components of products. In this context, it is important to note that for many categories of products covered by NLF legislation listed in Annex II, Section A, the involvement of a third-party in the conformity assessment procedure can be avoided by applying respective harmonised standards according to the options for conformity assessment laid down in the individual NLF legislations. It is therefore doubtful if this criterion in article 6 for classifying AI systems as high risk („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") is useful as a classification rule.[2]

## Risk classification under the Council's compromise

Art. 6 (3) sets out an exception to the risk classification according to Annex III. It stipulates that 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 therefore not likely to pose a significant risk to the health, safety, or fundamental rights. In order to ensure uniform implementation of the AIA, the EU Commission shall, no later than one year after the entry into force, adopt implementing acts to specify the circumstances where the output of AI systems referred to in Annex III would be purely accessory.

Regarding the classification of AI systems as high risk, the compromise proposal, thus, includes an additional horizontal layer on top of the high-risk classification made in Annex III, in order to ensure that AI systems that are not likely to cause serious fundamental rights violations or other significant risks are not captured.

## Risk classification under the EP's amendments

With regard to high-risk systems pursuant to Annex III, the EP's approach provides for an **additional assessment whether the AI system poses a significant risk.** According to art. 6 (2), 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 high-risk if it poses a significant risk of harm to the environment.

Moreover, the EP obliges the EU COM to provide guidelines at least 6 months prior to the entry into force of the AIA, following consultation with the AI Office and relevant stakeholders. These guidelines shall clearly specify 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

[2] At the same time, in case of medical devices the involvement of third-party conformity assessment bodies is mandatory for almost all of the device classes (3 out of 4 classes are subject to this independent assessment) wherewhile only medical devices of class III and in vitro diagnostic medical devices of class D pose the highest risk.

not. If a provider misclassifies their AI systems, they are subject to fines according to art. 71.

However, the EP text also contains a new mechanism for a pre-marketing notification to regulators of an AI system covered by Annex III they consider not to pose a significant risk as described and is therefore not covered. The competent authority is obliged to inform the provider within three months if they deem the AI system to be misclassified.

**Bitkom position**

We welcome both the Council's and the EP's directions of amendments. Both aim at ensuring that only AI systems that really carry a high risk have to fulfil the AIA's obligations for high-risk systems. While the Council's approach is a good start, we are even more in support of the EP's idea of "significant risk" ". In our view, this approach expresses more clearly the intention of the concept of  "purely accessory" used by the Council. Underlying "significant risk" with the definition of risk introduced by the New Legislative Framework (NLF) helps to set the basis for an effective risk-based regulatory approach and answers one of Bitkom's first demands[3]. The exact interpretation of "significant" will of course be highly relevant for the implementation of the regulation. Delaying this specification further in time creates prolonged legal uncertainty for companies. However, if the EU COM will then support the companies' classification of their AI systems by providing guidance, the operationalization of the legal text will be facilitated

However, we reject the EP's addition of an obligatory suspensory notification regime for Annex III systems that providers consider not to pose significant risk. As currently formulated, this obligation will be unfeasible in practice as the expected waiting time of (at least) three months will delay innovations and make them more expensive. It is also not clear what will happen if the competent authority does not respond within three months. In view of the expected additional workload the authorities are facing and the high penalties threatened under paragraph 2b, this is will not be conducive to Europe as an attractive location for innovation. Furthermore, the waiting time is counterproductive to the usual operating principles in agile project structures. However, making this notification regime voluntary could provide companies with guidance for unclear cases. Therefore, we propose to offer it as a voluntary option.

Overall, we would like to highlight that by linking the concept of high-risk to the severity of harm, the EP introduces a clear and established concept that is also flexible enough for supervisory authorities. We recommend that the co-legislators establish a definition of high-risk that provides for the necessary flexibility and legal certainty, as proposed by the EP .

With regard to high-risk systems covered by the NLF, it is imperative to revisit the principles for classifying the AI systems as high risk. The criteria for the assessment

[3] 2021august_bitkomposition_aiact.pdf

should be less formalistic and rather based on the risk that certain products may pose to health, safety or fundamental rights.

## General Principles and AI Literacy

In art. 4a, the EP has added general principles applicable to all AI systems. These principles entail human agency and oversight, technical robustness, privacy and data governance, transparency, diversity, non-discrimination and fairness, and social as well as environmental well-being. These principles are not binding but providers shall make their best efforts to comply with them.

In addition, the EP stipulates that both EU and Member States shall promote measures for the development of a sufficient level of AI literacy, across sectors and taking into account the different needs of groups of providers, deployers and affected persons concerned.

**Bitkom position**

The principles of AI use correspond to the values advocated in AI ethics and should be emphasised. The cross-sectoral promotion of AI knowledge is also important for the successful implementation of AI technologies. It is therefore to be hoped that the member states will take promotional measures.

# High-risk systems under Annex II

AI systems are considered high-risk if they are products or safety components of products regulated by the New Legislative Framework (NLF) legislation listed in Annex II, Section A, and according to which the product must undergo a third-party conformity assessment.

**Bitkom position**

With regard to these AI systems, it is highly important to ensure that the obligations under the AIA and under sectoral regulation ( NLF) do not overlap or conflict and do not create a double compliance burden. It needs to be ensured that the conformity assessment procedures that are available under the respective NLF legislation listed in Annex II, Section A, can be applied also to demonstrate conformity with applicable AIA requirements. Certificates issued according to NLF legislation listed in Annex II, Section A, shall integrate pertinent information related to the AIA conformity assessment, so that overhead and disruption of business caused for example by different certificate expiration dates or split conformity assessment procedures is avoided. Further, guidance on how to comply with AIA requirements in conjunction with applicable sectoral legislation listed in Annex II, Section A, shall be provided by the AI Office in cooperation with responsible sectoral bodies. This point also applies to and is further discussed in the chapter on “Interplay with existing legal frameworks”.

# High-risk systems under Annex III

## EU COM proposal

In Annex III, the EU COM proposal lists areas of application and subordinate use cases which are deemed high risk. The list entails the areas of biometric identification and categorisation of natural persons, management and operation critical infrastructure, education and vocational training, employment, workers management and access to self-employment, access to and enjoyment of essential private services and public services and benefits, law enforcement, migration, asylum and border control management, and administration of justice and democratic processes. In addition to listing the areas, Annex III then refers to specific use cases within these areas that are regulated as high-risk under the AIA.

## Council amendments

The list of high-risk AI use cases in Annex III has been amended by the Council. On the one hand, three use cases have been deleted (deep fake detection by law enforcement, crime analytics, verification of the authenticity of travel documents). On the other hand, two new use cases have been added (critical **digital** infrastructure and life and health insurance).

Moreover, art. 7(1) has been modified in order to provide for a possibility not only to add high-risk use cases to the list by means of delegated acts, but also to delete them. art. 7 (3) of the compromise text sets out the conditions under which the EU COM may remove AI systems from the list in Annex III.

## EP amendments

The EP adds critical digital infrastructure, systems for placing targeted job advertisements, systems for task allocation based on individual behaviour or personal traits or characteristics, systems that (help) decide on the eligibility for healthcare services and essential services, including but not limited to housing, electricity, heating/cooling and internet to the list of high-risk use cases.

**Bitkom position**

**Biometric Identification (BID):** The vast majority of AI in BID is not deployed in sensitive areas but enables routine provision of services that make life easier or that entertain. The EP proposes to expand the high-risk category to also include “biometric-based systems” and “AI systems intended to be used to make inferences about personal characteristics.” These terms would expand this category and include a broad range of use cases that may indirectly link to biometric data, including non-personal data. The proposed additions would significantly hamper this enabling technology. We therefore do not support the EP’s proposed extension to biometric-based data and AI that makes inferences. BID prohibition should be limited to specific biometric identification used by public entities. It should not include AI that requires interaction with the user and is therefore not intended for surveillance. This is consistent with the Council proposal. Instead, we support the exclusion of BID that serves authentication/ verification. The latter should also include systems where data of several individuals are compared against data in a data base.

**AI in the Workplace:** The vast majority of AI in the workplace does not cause harm but drives efficiency and ensures safety. For the most part, it focuses not on individuals but on general workplace processes. The monitoring and evaluation of performance is only to be considered high-risk if it is actually monitoring of employees. AI at the workplace may also be used to improve customer experience as “personal traits & characteristics” may use AI to match a customer service agent with a specific customer demand (e.g. considering language skills). Defining all these AI uses as high-risk would place compliance burdens on beneficial use cases. EP and Council and Commission all have almost the same text; but the exception of "significant risk of harm" and "accessory AI" would help here.

We recommend retaining the language proposed by the EP for paragraph 4 of Annex III, in particular the changes made to letter b) with the inclusion of “materially influence”, which should also be reflected in letter a) of the same paragraph. This would ensure consistency for deployment of AI in the employment sector, and the necessary legal certainty for AI providers and deployers in qualifying when an AI would fall within the scope of the AI Act.

Even this approach can limit essential functions of AI in the world of work, especially in recruiting and human resources. The general classification of AI as high-risk in the field is not appropriate. AI is already used for recruiting and human resources. The potential to solve many problems such as the shortage of skilled workers by using AI is enormous.

In addition, the classification does not recognise future developments and potential changes in the world of work. For example, it cannot be ruled out that automated job platforms and freelancing could become more of a focus in the future and that this classification stands in the way of development.

Moreover, there are already protections for individuals not to be subject to the decisions of automated processes under article 22 GDPR. However, the preparation of these decisions should still be possible by an AI system. This is particularly useful in companies with a large workforce.

**Recommender Systems**: The EP introduced a new high-risk classification for recommender systems used by social media very large online platforms (VLOPs). Recommender systems for user-generated content are defined in Point 8(ab) of Annex III as “AI systems intended to be used by 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.”

Recommender systems provided by VLOPs are already extensively regulated by the recently negotiated Digital Services Act (DSA): Providers of VLOPs are required under article 34 of the DSA to conduct annual risk assessments “diligently identify, analyse and assess any systemic risks in the Union stemming from the design or functioning of their service and its related systems, including algorithmic systems, or from the use made of their services” (emphasis added). These risk assessments cover a wide range of potential risks, including the dissemination of illegal content, negative effects on fundamental rights and mental well-being, the protection of minors and public health. Based on these risk assessments, article 35 of the DSA sets out that VLOPs must implement measures to mitigate risks, which includes the “testing and adapting their algorithmic systems, including their recommender systems”. VLOPs must also provide regulators and independent auditors access to recommender systems to allow them to understand how they are functioning. Auditors’ operational recommendations must be taken into account when taking necessary corrective actions. In addition to these measures to deal with any potential systemic risks recommender systems might pose, all online platforms are also obliged to provide transparency on the use and functioning of recommender systems. New regulatory initiatives should provide clear value by responding to hitherto unaddressed risks. Any new requirements should also be proportionate to the risks identified. By contrast, the introduction of additional, duplicative requirements as proposed by the EP creates legal uncertainty and imposes unnecessary and disproportionate regulatory burdens. Moreover, for the AIA, it is questionable that the high-risk classification should be linked to the size of the provider, i.e. whether or not the provider is a VLOP. It can be argued that a particular AI system (such as a recommender system) either presents certain risks or it does not and the fact that it is deployed by a company with more users rather than a smaller one should play no role.

To the extent there are any perceived risks associated by the VLOPs’ use of recommender systems in their services, these are already (and more appropriately) addressed as part of the broader systemic risk approach designed by the DSA.

We therefore disagree with the classification of these types of systems as high-risk AI and believe the provision should be removed in the final text of the AIA.

As such, there is a serious risk that introducing these additional obligations in the AIA will lead to further duplication of legal requirements under EU law, which will do nothing additional to address concerns but will only create additional confusion due to the overlapping rules. We are also worried that this duplication will later lead to confusion in the enforcement of rules, where different entities are involved.  We discuss this in detail in the section "Interplay with existing legal frameworks".

# AI definition

The definition of AI systems is one of the most controversially discussed parts of the proposed AIA. With good reason, because the scope, precision and discriminatory power of the entire regulatory framework depends on the definition of AI.

## EU COM Proposal

According to the EU COM proposal, AI is defined as a software which is developed with machine learning approaches, logic- and knowledge-based approaches, statistical approaches and search and optimization methods.

Since the aforementioned definition would entail a wide range of software used since decades, we strongly advocate for narrowing and sharpening the definition of AI. With regard to the high number of strict requirements set out by the proposed AIA, it is essential to only encompass systems which pose a new and different risks compared to conventional software.

The risk of including (advanced) software tools in general, must be eliminated during the trilogue negotiations. A shift away from the original definition by the EU COM is therefore indispensable.

## Council amendments

The Council  defines AI as a system that is 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. In order to make AI technologies easier to distinguish from classical software systems, the compromise text narrows down the definition in art. 3 (1) to systems developed through machines learning approaches and logic- and knowledge-based approaches. An explanation of the terms 'machine learning approaches' and 'logic- and knowledge-based approaches' is given in recitals 6a and 6b.

Machine learning is defined as a system's capacity of learning and inferring from data to solve an application problem without being explicitly programmed with a set of step-by-step instructions from input to output. The recital non-exhaustively enumerates different machine learning techniques (supervised, unsupervised, reinforcement learning) and architectures (neural networks, statistical techniques (e.g. logistic regression, Bayesian estimation)).

Logic- and knowledge-based approaches equip systems with logical reasoning capabilities on knowledge to solve an application problem. Such systems typically involve a knowledge base and an inference engine that generates outputs by reasoning on the knowledge base (expert systems). The recital ends with a non-exhaustive list of examples (knowledge representation, inductive (logic) programming, knowledge bases, inference, and deductive engines, (symbolic) reasoning, expert systems and search and optimisation methods).

Annex I and the corresponding empowerment for the EU COM to update it by means of delegated acts has been deleted. Art. 4 does not refer to Annex I any longer. Instead, it stipulates that the Commission may adopt implementing acts to specify the technical elements of machine learning, logic- and knowledge-based approaches, considering market and technological developments. Those implementing acts shall be adopted in accordance with the examination procedure referred to in art. 74(2).

## EP amendments

The EP has chosen a definition which is close to the OECD definition of AI. According to that, AI system means a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions that influence physical or virtual environments

**Bitkom position**

Both proposals incorporate the **notion of autonomy** into the definition of AI. We consider this a **useful addition** as it makes explicit where the potential risks of AI come especially into play: when there is no human control. However, exactly defining what "degree" or "element" of autonomy is enough to be considered AI is still open for interpretation. This will lead to legal uncertainty and needs to be addressed in further guidance.

We understand both proposals to narrow down the definition compared to the general interpretation of the EU COM proposal. We support especially the EP's highlight on machine learning and the reinforced reference to the indicative black box characteristic in the recitals. We urge that in trilogue this focus on what AI specific traits lead to additional risks is strengthened.

We would add that the EP put forward a definition of AI that is in line with the work of the OECD, and likely to become an internationally accepted definition of Artificial Intelligence. Most importantly, the AI definition should be aligned with existing international definitions, in order to provide companies with legal certainty.

# Prohibited AI Practices

**Biometric Identification (BID)** (art. 5): The prohibition of AI used for biometric identification (BID) is an area of divergence between  Council and EP. The EP's proposed blanket ban on BID in the public to tackle risks of mass surveillance would outlaw beneficial use cases and risks hampering this enabling technology. The EP's position extends this to systems based  on or inferred from biometric-based data. This wide scope is problematic, as it would negatively effect many products which are clearly beneficial and not harmful to users. Many products provided by private entities use AI for biometric identification (e.g., voice assistants), and biometric identification can be useful in health-related applications.  These systems are not used for surveillance and do not cause harm[4]. This includes, for example, products to provide access to people with disabilities and routine provisions of personalised services. The prohibition should be limited to specific biometric identification used by public entities and exclude AI that requires interaction with the user and is therefore not intended for surveillance.

The EP also proposes a ban of biometric categorisation systems that deal with sensitive personal characteristics regardless of the intended use of these systems or whether users have given their consent (art 5.1ba). If adopted, this would have serious unintended negative effects on legitimate business practices and innovation, particularly in the context of fairness training and bias correction for content

[4] While it is hard to argue that some system is completely incapable of causing harm, many systems only hold a vanishingly small risk.

moderation tools on social media platforms, to ensure that error rates for detecting harmful content are consistent across protected classes, as well as immersive technologies at a time when the EU is trying to seize the opportunities of these technologies through a separate "virtual worlds" initiative. For example, these low-risk biometric AI systems allow consumers to virtually try on clothes and accessories before buying them online, helping to reduce environmentally harmful product returns; give consumers the experience of creating personal avatars for use in messaging or gaming; and enable the virtualisation of cartoon characters or animals on people's real shoulders.

Bitkom position

With regard to the principle of proportionality, prohibitions should be strictly limited to specific use cases. We therefore support the EU COM and the Council positions that low risk and beneficial biometric categorisation systems should  only be subject to the transparency obligations under art. 52, and encourage the co-legislators to remove the ban proposed by the EP.

# GPAI, Foundation models & Generative AI

GPAI**,** foundation models, and generative AI are AI systems that can be used in  a wide range of possible applications and use cases, both intended and unintended by their developers. They can be applied to many different tasks in various fields, often without substantial modification and fine-tuning. They are characterized by their scale) as well as their reliance on transfer learning (applying knowledge from one task to another).

GPAI encompasses pre-trained models which can be used for more specialised AI systems (e.g. natural language processing systems which can be used for chatbots, decision assistants, translation etc.) When deployed  by third parties, specific risks stemming from the use of a GPAI system relevant to that application or use  are therefore often impossible to predict and mitigate by the developers of GPAI. Foundation models are trained on broad data at scale and designed for generality of output, while generative AI (a subset of foundation models) is specifically intended to generate content.

## EU COM proposal

In the EU COM proposal, GPAI is not mentioned explicitly. Only recital 60 refers to the complexity of the AI value chain and provides for suppliers of pre-trained models to cooperate, as appropriate, with providers and users to enable their compliance with the obligations under the AIA.

## Council amendments

The Council, on the other hand, mentions GPAI explicitly. Art. 3 (1b) defines GPAI as a system that - irrespective of how it is placed on the market or put into service, including as open-source software - is intended by the provider to perform generally applicable functions such as image and speech recognition, audio and video generation, pattern detection, question answering, translation and others. A GPAI system may be used in a plurality of contexts and be integrated in a plurality of other AI systems.

Art. 4b (1) obliges providers of GPAI systems which may be used in high-risk systems to adhere to the provider obligations set out in art. 8-15. However, only to an extent laid down by an implementing act which is based on prior consultation and detailed impact assessment.  According to art. 4c, providers can explicitly exclude all high-risk uses in the instructions of use and, as a result, liberate themselves from adhering to the provider obligations set out in art. 8-15.

## EP amendments

The EP distinguishes between foundation models, GPAI, and generative AI. According to art. 3 (1c), foundation models are AI systems capable of producing general output and designed as templates for specific applications. In line with this logic, art. 3 (1d) and Recital 60e refer to GPAI as usable applications of foundation models that can fulfil a large number of unanticipated tasks. Finally, according to art. 28b (4), generative AI is a specific application of a foundation model intentionally designed for creating content fully or partly autonomously.

The EP focuses on foundation models. Art. 28b sets out obligations for providers of foundation models, entailing constant analysis, monitoring and reduction of risks to "health, safety, fundamental rights, the environment and democracy and the rule of law", taking appropriate data governance measures for the mitigation of possible biases, a lifecycle-long evaluation system for ensuring "appropriate levels of performance, predictability, interpretability, corrigibility, safety and cybersecurity". Moreover, developers of foundation models must reduce the use of energy and resources, enable the measurement and logging of the consumption of energy and resources, provide technical documentation and instructions for downstream providers, register the model in an EU database, and introduce a quality management system. In addition to that, they must keep the technical documentation at the disposal of the national authorities for a period ending 10 years after their foundation models have been placed on the market or put into service.

In case the foundation model is designed for or used in generative AI applications, the provider of the model must also develop mechanisms that prevent the creation of illegal content and publish a "sufficiently detailed summary" of the copyright-protected training data. Moreover, providers must comply with the transparency obligations set out in art. 52.

These obligations apply regardless of whether the foundation model 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.

### Bitkom position

We observe that under the current impressions of what AI systems can be used for, a lot of new aspects entered the discussion of the AIA. In general, we deem it critical to deviate from the risk-based approach by including GPAI and foundation models explicitly into the act. To our understanding, GPAI systems were already taken into account in the EU COM proposal: On the one hand in recital 60, setting out cooperation obligations, on the other hand, as soon as these systems are used for high-risk purposes and art. 8-15 apply and GPAI providers must make sure that any used underlying model can deliver the necessary information.[5]

When integrating GPAI, foundation models or Generative AI into the legislation, our proposal, thus, is to strengthen this mechanism of information sharing within the value chain. This helps to accommodate various possible AI-component-compositions and makes sure that the additional requirements only affect systems that are used in high-risk settings. This ensures a proportionate approach which does not create burdens where no significant risk is given.

The EP proposal goes in that direction by providing more explicit rules for cooperation within the value chain. However, the far-reaching obligations defined for foundation models in art. 28b undermine the concept of value chain cooperation by shifting most responsibilities to the providers of such models and being partly unfeasible. Since the idea of foundation models is to have a general template for countless unforeseen applications, it is extremely costly and sometimes impossible for foundation model providers to anticipate and continuously analyze risks, environmental impact, and illegal content produced by the models. This one-sided focus on foundation model providers will seriously constrain the development and use of such models in Europe. At the same time, downstream deployers using foundation models as a basis for their applications need insights and control mechanisms for the models to fulfil their responsibility within the value chain. The burden must not be shifted on deployers who neither have data, know how to fulfill requirements but there must be clear distribution of responsibilities and cooperation mechanisms in place.

Therefore, we propose to **remove the monitoring of risks, mitigation of possible biases, evaluation system, reduction of energy and resources, and prevention of illegal content creation requirements**. Instead, a stronger focus should be set on technical documentations and instructions and a quality management system provided by the developer of a foundation model to adequately help downstream providers fulfil their

[5] Overall, it should be acknowledged that any targeted requirements for foundation models could lead to unintended effects on all levels of the value chain. This risk is also seen by numerous other stakeholders, such as more than 150 European users of AI, who spoke out on the issue in an open letter at the end of June, urging EU institutions to return to a risk-based approach to AI regulation in the AI Act, without specific requirements for foundation models (Source Heise: Link).

obligations according to the risk-setting of their applications. However, such documentations, instructions, and quality management systems need to be designed in a way that they safeguard trade secrets. Similarly, the obligation to publish a summary of copyright-protected training data is unfeasible given the amount of data and number of different copyright legislations worldwide. Copyright-specific solutions should be defined separately from the AI Act, if required. This should consider existing mechanisms on text and data mining included in the EU Copyright Directive and jurisprudence such as on search engines and indexes. This would avoid overburdening the AI-specific negotiations.

Furthermore, it is essential that the final AIA provides **clear definitions of foundation models, GPAI and generative AI**, **and specifies their relationships**. In the current proposals by the Council and EP, these concepts are partly used with different meanings and in different contexts, which causes uncertainty and confusion. In any case, it is essential that providers of foundation models, GPAI, or generative AI are obliged to support the deployer in a reasonable way to ensure compliance whenever required.

Moreover, the final legislation should entail exceptions for foundational models which are not passed on to a third party but are only deployed internally by the provider.

Finally, as long as in conformity with the regulations enshrined in the AI Act, business parties should be free to allocate responsibilities through contractual obligations for GPAI, foundation models, and generative AI.

# High risk systems - provider obligations

The AIA will set out a number of obligations that must be fulfilled by providers of high-risk AI systems. These rules entail ex ante market access obligations, which must be fulfilled before high-risk AI systems can be placed on the market, as well as ongoing ex post market monitoring obligations.

## Council amendments

Many of the requirements for high-risk AI systems, as provided in Chapter 2 of Title III of the EU COM proposal, have been amended by the Council in order to make compliance for providers of high-risk systems easier.  In specific, the following amendments have been made:

In art. 8 (compliance with requirements), compliance obligations are limited to the generally acknowledged state of the art.

Art. 9 (risk management system) has been amended as follows: Art. 9 sec. 2b has been deleted. An estimation and evaluation of the risks that may emerge when the high-risk AI system is used in accordance with its intended purpose and under conditions of reasonably foreseeable misuse is not necessary any longer. The following has been added to art. 9 sec 2: The risks referred to in this paragraph shall concern only those which may be reasonably mitigated or eliminated through the development or design of the high-risk AI system, or the provision of adequate technical information.

Art. 10 (data governance) has been specified: Art. 10 sec. 2f is now limited to biases "that are likely to affect health and safety of natural persons or lead to discrimination prohibited by Union law" instead of biases in general.

In art. 11(1) (technical documentation) the Council added an exception for SME's (including start-ups). In the case of SMEs, including start-ups, any equivalent documentation to the documentation according to art. 11, meeting the same objectives, unless deemed inappropriate by the competent authority is sufficient.

## EP amendments

With regard to the obligations for high-risk AI systems, the EP addresses deployers in addition to providers. Their obligations are listed in art. 29.

According to the EP, providers of high-risk AI systems must, inter alia, register themselves, inform the natural persons who function as human oversight about automation bias, comply with accessibility requirements, and

> *provide specifications for the input data, or any other relevant information in terms of the datasets used, including their limitation and assumptions, taking into account the intended purpose and the foreseeable and reasonably foreseeable misuses of the AI system* (art. 16 (ac)).

Art. 18 (technical documentation) has been deleted. The conformity assessment has been moved from art. 19 to art. 33 ff..

In case corrective actions become necessary, providers are obliged to inform the distributors, the importers, the national competent authorities of the Member States in which they made the AI system available or put it into service and the deployer, where possible.

According to art. 23, upon reasoned request providers must grant national competent authorities with access to the logs automatically generated by the high-risk AI system, to the extent such logs are under their control. The information shall be considered a trade secret.

The EP added a limitation to art. 10 (data and data governance), stating that data quality must only guaranteed as far as this is technically feasible according to the specific market segment or scope of application. In addition, data quality requirements for unsupervised learning methods have been addressed specifically in art. 10 (2).

Where the provider cannot comply with the obligations laid down in this art. 10 because it does not have access to the data and the data is held exclusively by the deployer, the deployer may, based on a contract, be made responsible for any infringement.

**Bitkom position**

The changes by the Council and the EP are to be welcomed. The original monitoring and evaluation obligations were excessive. The current changes are more oriented towards technical feasibility and set more realistic conditions. As already noted, the demands on the deployer must be lowered. Above all, the extreme monitoring and reporting obligations are a disproportionately high bureaucratic hurdle. Furthermore, the obligation to consultation with worker's representatives might lead to hesitation and less implementation of AI.

Therefore, we welcome the Council proposal excluding AI systems whose output is “purely accessory” from “high-risk.” Also, we support the EP’s proposal limiting high-risk classification to AI that imposes significant risk of harm to health, safety, and fundamental rights. “Significant harm” should be specified in a way that excludes AI that does not create a material adverse risk.

# Fundamental Rights Impact Assessment

## EP amendments

The Fundamental Rights Impact Assessment (“FRIA”) has been introduced by the EP. The obligation to conduct a FRIA prior to putting into use a high-risk AI system is directed towards deployers of AI systems and entails questions about the purpose, scope, lawfulness, and the system’s foreseeable impact on fundamental rights. The FRIA must be conducted before the system is put on the market and relevant stakeholder must have the opportunity to give their feedback in advance.

**Bitkom position**

The EU COM as well as the Council have already had the protection of fundamental rights in mind and have provided for it sufficiently, for example through the provisions in articles 8-15 (especially article 10). In addition to that, under the EU COM's article 29 users must monitor the risk of the AI-system during its operation. By introducing an additional fundamental rights impact assessment, we see a double regulation. While fundamental rights are undoubtedly worth protecting, we question whether the FRIA will act as an additional safeguard. It rather makes users confirm the results of AI providers – at the cost of a no-value-adding bureaucracy.

# Roles and responsibilities

## EU COM proposal

The EU COM proposal is directed mainly at two actors: providers and users of AI systems. The provider's obligations are listed in art. 8-23 while the user's obligations can be found in art. 29. In addition to that, obligations of manufacturers, authorised representatives, importers, and distributors are stipulated in art. 24-27. Art. 28 provides rules concerning the transition from user to provider.

## Council amendments

Since AI systems are developed and distributed through complex value chains, the Council's compromise text includes changes clarifying the allocation of responsibilities and roles. With regard to art. 13 and 14 the compromise text has added provisions that allow for more flexible cooperation between providers and users. Furthermore, art. 23a indicates more clearly the situations in which other actors in the value chain are obliged to take on the responsibilities of a provider.

## EP amendments

According to the EP, art. 28 sets out responsibilities along the AI value chain of providers, distributors, importers, deployers or other third parties.

A deployer shall be treated as a provider, if they make a substantial modification to an AI system, including a GPAI system, which has not been classified as high-risk and has already been placed on the market or put into service in such manner that the AI system becomes a high risk AI system in accordance with art. 6.  In addition, the former provider shall provide the new provider with the technical documentation and all other relevant and reasonably expected information capabilities of the AI system, technical access or other assistance based on the generally acknowledged state of the art that are required for the fulfilment of the obligations set out in the AIA.

Also, third parties that supply tools services, components or processes that are used or integrated in the high-risk AI system shall provide information to enable the provider of the high risk AI system to fully comply with the AIA. For that purpose, the EU COM shall draft non-binding contractual clauses.

Art, 28(a) stipulates that unfair contractual terms unilaterally imposed on an SME or start-up shall not be binding.

**Bitkom position**

The mechanism proposed by the EP with regards to the former provider being responsible of supplying the necessary information to the next provider, comes close to what we imagine the adequate handling of the complex value chain for AI systems. While the final provider/deployer is best suited to judge the threats of the AI system, she might require certain information from actors up the value chain. Thus, cooperation with the final goal of risk-minimisation should me supported by the AIAt.

# Harmonised standards

## EU COM proposal

Harmonised standards as defined in Regulation 1025/20126 according to art. 40 are key to show compliance with the provider obligations. If providers of AI systems adhere to harmonised standards, it is rebuttably presumed that the AI systems are placed on the market in compliance with the AIA. We expressly welcome this approach.  The use of common specifications according to article 41, on the other hand, should only take place in absolute and justified exceptional cases when safety or fundamental rights are not properly addressed in the standards requested by the EU COM.  Therefore, one of the two central fields of action for the innovation-friendly implementation of the AIA is an active-strategic design of the landscape of horizontal and vertical standards that enable proof of compliance with the respective requirements.

## Council amendments

Art. 40 (1) on harmonised standards has been modified slightly by adding GPAI systems. In addition to that, the compromise text adds art. 40 (2) which makes specific provisions regarding standardisation processes. It refers to the procedures laid down in art. 10 Regulation 1025/2012[7] and lays down a non-exhaustive list of objectives which must be taken into consideration by the EU Commission within the standardisation process. These objectives comprise:

> a) ensuring that AI systems placed on the market or put into service in the Union are safe and respect Union values and strengthen the Union's open strategic autonomy;
>
> b) promoting investment and innovation in AI, including through increasing legal certainty, as well as competitiveness and growth of the Union market;

[6] According to Regulation 1025/2012 article 2 paragraph 1c a „harmonized standard” is a „ a European standard adopted on the basis of a request made by the Commission for the application of Union harmonisation legislation”.

[7] Regulation 1025/2012 has been amended in 2022 and will be evaluated starting in Q3/2023.

c) enhancing multistakeholder governance, representative of all relevant European stakeholders (e.g., industry, SMEs, civil society, researchers);

d) contributing to strengthening global cooperation on standardisation in the field of AI that is consistent with Union values and interests.

Moreover, the standardisation organisations shall provide evidence of their best efforts to fulfil these objectives.

With regard to art. 41 the compromise text limits the EU COM's discretion regarding the adoption of implementing acts establishing common technical specifications for the requirements for high-risk AI systems and GPAI.

## EP amendments

The presumption of conformity laid down in art. 40 (1) has been expanded to foundational models and to the obligations that their providers must fulfil.

In addition, the EP aims at setting a deadline of two months for the EU COM to issue standardisation requests after the AIA entered into force. Similar to the Council's approach, the EP intends to make more precise provisions on regarding the standardisation process. On the one hand, the EP focuses on coherence with the standards that apply to the NLF already, on the other hand, innovation, investments, competitiveness, and a balanced consideration of all relevant interests are to be ensured.

According to art. 41 of the EP's approach, the EU COM may only adopt common specifications if there are no harmonised standards, and a prior standardisation request, without reason, has not been accepted by any of the European standardisation organisations.

**Bitkom position**

We welcome the proposal of both Council and EP to introduce certain requirements that need to be fulfilled before the EU COM may resort to common specifications. Priority should always be given to standards developed within the European standardisation system with the involvement of industry stakeholders to allow effective implementation and coherence within the body of standards. Only when the listed measures are exhausted, common specifications should be considered an option. Regarding proposals that give certain guideline of how standards should be developed, we recommend leaving this to the respective standardisation regulation (1025/2021) as these principles should be consistent over all areas where standardisation is used. We welcome the EP's intention of assuring the timely development of standards by giving a deadline to the EU COM until when the standardisation requests need to be issued. However, besides the requests, also the timeline with regards to the entering into force of the regulation is relevant here.

# Regulatory sandboxes

## EU COM proposal

The EU COM proposal provided for the establishment of regulatory sandboxes in art. 53. These sandboxes shall provide for a controlled environment that facilitates the development, testing and validation of innovative AI systems for a limited time before their placement on the market or putting into service pursuant to a specific plan.

## Council amendments

In order to create a legal framework which is more innovation-friendly, art. 53 has been changed substantially. Amongst other things t has  been clarified that AI regulatory sandboxes, which are supposed to establish a controlled environment for the development, testing and validation of innovative AI systems under the direct supervision and guidance by the national competent authorities, should also allow for testing of innovative AI systems in real world conditions. Furthermore, new provisions in art. 54a and 54b have been added allowing unsupervised real-world testing of AI systems, under specific conditions and safeguards.

## EP amendments

According to the EP, member states shall establish at least one regulatory sandbox at national level which shall be operational at the latest on the day of the entry into force of the AIA. Establishing authorities are obliged to allocate sufficient resources to comply with this provision in timely manner.

### Bitkom position

**The AI Act should not only regulate AI, but also promote its development and implementation in the European Union**: for an ecosystem not only of trust but also excellence. Sandboxes are one option to foster this. It must be ensured that they are designed in a way that really provides additional value to companies using them.

# Interplay with existing legal frameworks

A smooth and well-functioning interplay between the AIA and existing regulation that applies to AI systems is extremely important for developers and distributors of AI systems as well as their users. A coherent legal framework contributes to more legal clarity which, in turn, is the basis for trust and investment in AI.

We welcome that the EU COM proposal is linked to existing horizontal and vertical regulatory frameworks. Explicit reference to the New Legislative Framework ("NLF") is helpful since the NLF has been established several years ago at the horizontal level and is since then well-known and adopted in practice. Since its adoption in 2008, the NLF has demonstrated its ability to support future proof legislation by making use of technical standards in order to substantiate indefinite legal concepts.

In addition, the AIA refers to a number of sector-specific EU regulations and directives. From a practical viewpoint it, it is essential that providers of products which are regulated by existing EU law do not face double requirements from both sector-specific regulation and the AIA.[8] During the trilogue negotiations, a main focus should be put on how to make market access and ongoing market monitoring feasible, especially for providers of high-risk AI systems.

The same applies for upcoming legal framework, such as the AI Liability Directive ("AILD") or the Product Liability Directive ("PLD"), which will refer directly to the AIA's provisions and will be closely linked systematically and in terms of content.

Policy makers regularly emphasise the overarching strategic goal of their AI policy: The creation of a European ecosystem of excellence in AI that is closely linked to an environment of trust in the use of AI. This should be the benchmark for the upcoming trilogue negotiations.

There is a serious risk that introducing these additional obligations in the AI Act will lead to further duplication of legal requirements under EU law, which will do nothing additional to address concerns but will only create additional confusion due to the overlapping rules. We are also worried that this duplication will later lead to confusion in the enforcement of rules, where different entities are involved. To make the risk of double regulation visible, we have compiled some concrete examples of possible duplication of efforts under the DSA:

- Article 9 risk assessment AI Act vs DSA risk assessment articles 34/35 (since this applies to our AI-powered recommender systems and content moderation tools).
- Transparency and provision of information to users in AI art 13 could have some potential overlaps with DSA articles 27 and 40.3. AI Act art 13 refers to explaining the characteristics, capabilities and limitation of performance of a high risk AI system. Art 27 DSA requires platforms to publicly explain the main parameters used in their recommender systems. Art 40.3 DSA requires a VLOP to explain the design, function and testing of the algorithmic system of a VLOP when requested.

[8] Medical devices and in vitro diagnostic medical devices, regulated under EU MDR and IVDR respectively, should be treated under the sectoral legislation and AI-related requirements can be added to MDR and IVDR in an upcoming revision of both regulations. In the AI Act, MDR and IVDR should be moved from Annex II, Section A to Section B. Due to the issues known already today with getting medical devices, especially innovative ones, released on the EU market under MDR/IVDR, any additional complexity with getting AI-based medical devices CE-marked needs to be avoided.

- On AI Act article 13 on transparency, references to levels of accuracy for AI systems vs reporting on content moderation practices per article 15 of DSA. Similar point on reporting on accuracy in article 15 of AI Act, vs article 15 of DSA on transparency on content moderation practices.
- AI Act article 13 obligation to report on known or foreseeable circumstances which may lead to fundamental right risks - could correspond to DSA article 34.1(b) re risk assessment impact on actual or foreseeable negative effects for fundamental rights
- AI Act article 13 obligation to specify characteristics of high risk AI systems relating to its performance as regards the persons or group of persons on which the AI system is intended to be used, has potential overlap with DSA article 27 on recommender system transparency. Similar overlap with DSA article 27 when it comes to AI art. 13 obligation to provide information on training data used.