

Remarks of the Polish Confederation Lewiatan on the draft of the Artificial Intelligence Act

The Polish Confederation Lewiatan expresses its deep satisfaction having learnt about a well-balanced approach of the European Commission to the direction of the artificial intelligence development within the European Union and appreciates the enumerative list of implementations for high risk applications as well as the fact of relying on internal conformity assessments along with industry-related codes of conduct. However, we have noticed the need to focus, during the subsequent legislative works on **ensuring competitiveness of enterprises in the European Union** in the field of artificial intelligence. We are hereby presenting the following postulates to current draft of the Act.

1. Too extensive definition of the artificial intelligence

Too extensive approach means that the scope of the regulation may cover both the systems commonly considered artificial intelligence as well as the systems operating in an only similar but less complex manner vis-a-vis artificial intelligence. Therefore, we suggest that simple, human-controlled machine learning be expressly left beyond the scope of the proposed regulation. **Not every single IT automation should be considered artificial intelligence as it is not composed of the predictive analysis automation but exclusively of some basic algorithms.** The future regulation on artificial intelligence should be simultaneously aligned with the machinery directive in respect of safety of persons, not safety of property. Moreover, in accordance with the theme no. 6 of the draft Act, the definition of artificial intelligence should be clear in order to ensure the **legal certainty while providing flexibility** to accommodate future technological developments. The Polish Confederation Lewiatan is of the opinion that artificial intelligence should be defined as the system performing analyses **with a dose of autonomy**. While in the current wording of the definition, the reference to the term of “autonomy” is missing. In Annex 1, the definition of artificial intelligence (para. 3 item 1) has been associated with the techniques and approaches in the field of artificial intelligence referred to in the relevant cases. Indeed, the Act states (para. 4) that the said list should be updated on a regular basis but this type of legislative technique does not correspond with the progress of civilisation in the field of artificial intelligence at all. Therefore, the adopted solution thoroughly misses the assumptions indicated in item 6 of the Preamble whereunder the term of artificial intelligence must be “*clearly defined to ensure legal certainty, while providing the flexibility to accommodate future technological developments.*” In this regard, a better definition has been proposed by the High Level Expert Group on Artificial Intelligence:

“Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals”.



2. Indication of specific limits of the prohibition of individual practices in the field of artificial intelligence

Due to the fact that some artificial intelligence systems are fully prohibited, it is exceptionally important to specify the limits of such a prohibition in a precise manner. This criterion has not been satisfied by para. 5 item 1 letter a) which prohibits to use the systems that are to impact recipients subliminally. Understanding this term may result in disputes which, in the end, may diminish innovativeness. Still, on the other hand it should be assumed that every single Member State has relevant protective laws in place, including those protecting consumers when it comes to subliminal message. Another issue is the term of psychological harm. In this regard the term “harm” should be rather understood as a physical harm. The content of item 16 of the Preamble does not eliminate doubts in this regard.

Therefore, we are in favour of formulating the definition of subliminal techniques and precise clarification of the “psychological harm” term.

3. Reference to fundamental rights

The reference to the term of “fundamental rights” in para. 7 item 1 letter b) is going to be the source of serious interpretational problems due to the bluntness of this term as well as the capacity for unlimited extensive interpretation. Adoption of the comprehensive catalogue of fundamental rights in its contemporary understanding would be tantamount to, for example, the necessity to implement a very broad perspective. Having applied only the most fundamental sources in this regard, such as the Charter of Fundamental Rights of the European Union or the European Convention on the Protection of Human Rights and Fundamental Freedoms we conclude that every single type of fundamental rights - both dignity and, for example, artistic freedom - should be taken into account. Along with each right the number of possible configurations increases which multiplies the risk - even up to the level that is hard to assess actually. **In such a situation, the legal certainty suffers from a noticeable detriment, leaving along the very fact that in practice a threat to human dignity causes completely different effects than limiting their artistic freedom.**

Covering the rights of all generations with the abovementioned term by default with simultaneous disputable nature of the so-called 4th generation rights (e.g. rights and freedoms related to human sexuality) precludes, in practice, the possibility for real and predictable estimation of the directions for possible changes in Annex 3 which translates into some **material decrease of the feeling of legal certainty.**

Another element which should be taken account of in this regard is the actual interpretative discrepancy of the fundamental rights term at the level of the Member States. This is an additional circumstance that would challenge legal certainty and generate the problem of its universality and uniformity of implementation. At this point we are also raising the issues of constitutional nature as in matters



pertaining to fundamental rights a strong and competitive-against-supradomestic-institutions role is played by courts and tribunals of constitutional nature which reserve their right of priority for the assessment of issues related to the protection of fundamental rights. A good example is the Republic of Germany where from the *Solange I* ruling of 1974 through subsequent rulings pertaining to constitutionality of the Community treaties to the ruling of 2010 on *Honeywell*, the German Bundesverfassungsgericht was persistently emphasising its superior role for matters related to fundamental rights.

The abovementioned term should be made materially precise by virtue of specification of the provision text or adding a theme which would narrow down directions for interpretation.

4. Possibility for data training even for some high risk applications

We support the concept assuming that for some internal procedures of artificial intelligence system efficient acts be taken up in order to counteract potential threats and create prudential framework for conducting internal declaration of conformity. A good example is the assessment of creditworthiness. A majority of artificial intelligence applications for the assessment of creditworthiness is managed internally with no need to assess the conformity externally and register in the databases managed externally. This process turned out to be relatively resistant to market turbulence. In the meantime, delegation of the conformity assessment for such implementations of artificial intelligence to an external regulator may result in a “bottleneck” and significantly **slow down the creation of new enhancements for models used in artificial intelligence systems and, as a consequence, create a precipice between EU markets and external ones.**

What is more, **testing some artificial intelligence systems, such as credit risk assessment, for example, to a lesser extent, with no need to treat them as high risk ones,** may still be favourable for consumers with simultaneous limiting the risk related to artificial intelligence. This is a common practice allowing for the selection of the best solutions relying on artificial intelligence available for concrete implementations.

Let us also draw your attention to imprecise and general rules of classification of artificial intelligence systems referred to in the Appendix 3 in conjunction with para. 6 of the draft Act. The proposed classification may result in covering even the simplest programs used with the term of high-risk artificial intelligence systems. For example, in the case of item 4 pertaining to the employment, staff management and access to self-employment, a description of artificial intelligence systems which should be deemed high risk systems is so general that virtually all systems used in the recruitment process may be deemed high-risk artificial intelligence systems.

We welcome the risk-based approach. However, the risk level should be determined by the criticality of the application itself and not by a sector approach. The criteria proposed to be considered are:

- analysis of unwanted consequences
- side effects

- administrative burden

We advise restricting the Annex III referring to high risk applications on education and employment only to situations without a final human decision making.

5. Conformity control

Para. 64 of the draft Act ordering to disclose and provide the data and documents requires explicit specification in terms of relationship against the business secret. As this information may be of confidential nature and frequently constitute the main, if not the only, element of the competitive advantage for a given solution. We are of the opinion that the rights given to domestic supervisory authorities over the rights market, such as **demand for access to the data collection, API interfaces and source codes are too far-fetched**. In particular, the absence of precise definitions for key risks (such as discrimination, partiality) does not increase the objectivity of supervisory assessment.

Additionally, **at the domestic level, actual procedural guarantees protecting business secret both from unauthorised access and excessive or unjustified access of individuals or entities to this type of information should be implemented**.

To sum up, the approach based on **the proportionality rule** should be included in the conformity systems for high-risk artificial intelligence applications.

6. Making definitions precise

The definitions should be clarified in order to make the assessment by supervisory authorities and users themselves objective.

The Polish Confederation Lewiatan is of the opinion that it is obvious that every individual should have access to the information related to the methodology for credit risk assessment applied to them. Defining “partiality” with regard to the creditworthiness assessment is of key importance so that a customer understands the outcome of the decision correctly even if it is made by artificial intelligence (unfair treatment may be understood in different ways). The decisions on creditworthiness rely on actual social and economic analyses, thus unequal treatment of social groups based on their features may be explained by presenting the data relying on statistics.

Since a notion of risk is a key component of the regulation, a specific definition of this concept that have to be considered by the risk management system must be added. While identifying risk management, it is advisable to make use of existing quality management systems (e.g. ISO 9001) instead of setting up a dedicated new AI quality management system.



The terms “bias” and “discriminatory effect” should be made more specific. A possible solution would be a “positive exclusion”, that is, indication when bias is admissible. Bias (para. 10 section letter f), theme 33) may be also defined as a “discrimination understood as a statistical error (from-above assigning features not consistent with reliable statistics) or from-above implementation of assumptions harmful to an individual”.

The suggested terminology, especially when it comes to bias, is potentially linked with a broadly understood anti-discriminatory issues, thus the issues of very extensive interpretative dynamics, diversity of opinions and regulatory diversity. What is more, the “bias” terminology suggests the need to avoid any forms of discrimination. Even those that are legally permitted both at the EU and domestic level.

We also suggest that the **term “child”** (within the framework of risk management system - para. 9 section 9) be made more specific. Firstly, indication of a concrete age limit would solve the problem of discrepancy at the level of domestic acts. Secondly, the assessment of impact on a child as part of the risk management system should rely on concrete age limits due to significant cognitive and emotional differences of children aged 7, 10 or 13. It is not possible to apply efficient criteria for all age groups. A good direction would be to specify what kind of impact on children is especially undesired as far as the regulation objectives are concerned.

7. Legal certainty and predictability of sanctions

Punishments should be imposed based on a clear catalogue of pre-requisites and concretely listed breaches. The Polish Confederation Lewiatan is of the opinion that the sanctions envisaged are too strict (especially for low-margin business entities) and **unproportionally increase business risk** related to the use and development of artificial intelligence. This will have a negative impact of the development of artificial intelligence in the EU. There is a risk that authorities will act arbitrarily with no specific administrative order. We suggest that clear mitigating and aggravating pre-requisites and circumstances resembling the ones available in the GDPR be implemented.

8. Obligations pertaining to transparency with regard to the specific artificial intelligence systems

Para. 52 imposes the obligation to inform a user that they enter into interaction with the artificial intelligence system unless it is obvious on the basis of the overall circumstances. The European Commission gives the example of a chatbot which would clarify this obligation based on an example. However, the language in its current wording is too vague, taking account of the fact that artificial intelligence has been integrated with numerous systems addressed to the user and used in order to obtain recommendations, searching for information, providing hints and forecasts. Since the definition of artificial intelligence system itself (para. 3 item 1) was based precisely on establishing, by artificial intelligence system, of the interaction with the user, then, in fact, every artificial intelligence system is going to satisfy this pre-requisite as it is a part of the very definition of artificial intelligence. Taking

account of the progress of artificial intelligence systems, there remains the issue of continued work on defining “interaction with natural persons”.

On the other hand, para. 52 section 3 imposes the obligation to inform on changing, by the user of AI, the objective reality, labelling this action as “deep fake”. Using the phrase “deep fake” in this context does not seem justified as this term suggests intentional misinformation of the recipient with regard to the identity of the presented object, frequently due to foul or even illegal motives. Obviously, such practices should be stigmatised. Making changes to the presented reality may, however, have a pole-like motivation - artistic, useful, explanatory, educational, quotative or polemical. At the same time it should be assumed that these are the situations constituting a majority of cases consisting in “manipulating” with the real picture. Imposition, in every case, of the obligation to inform on the change introduced even to the least extent, could materially hinder running journalist, artistic or, broadly speaking, artistic activity. The second section of the item referred to above eliminates this risk to a lesser extent. The legislator should specify whether every change should be labelled.

9. Disturbed balance of obligations among suppliers, implementing parties and users of high-risk artificial intelligence

In their current wording, laws make no difference between the obligations imposed on the AI user if they perform the role of a party implementing a given use of artificial intelligence system and the obligations of “artificial intelligence system supplier” towards the customer.

The party implementing the use of artificial intelligence should be ultimately a principal subject of the assessment as enterprises offering tools based on artificial intelligence systems are not able, in the end, to verify final use to which their systems are applied or additional data which may be introduced into the system. Suppliers of solutions relying on artificial intelligence systems may and should deliver any and all information that is necessary for the implementing parties to perform self-assessment. It is very important for a supplier of API solutions/ interfaces of which the supplier of solutions relying on artificial intelligence systems does not take over control when users give their consent to allowing customers/ users for access to the solutions at their discretion.

10. Exemptions pertaining to multi-task systems/ open source tools

The obligations to observe requirements pertaining to artificial intelligence systems should be vested in legal entities or natural persons using the tools of open source type, such as TensorFlow or AutoML as they have a final control over their objective and using artificial intelligence implementations. Imposition of obligations on open source tools suppliers would, to a large extent, discourage from providing such technologies which support the entire eco-system of innovativeness.

We are also in favour of exempting from obligation to publish fundamental research as the publication of fundamental research is not qualified as “placing on the market” or “handing in for use”.



The Polish Confederation Lewiatan points out that the required clarifications/safeguards, such as at least guaranteeing error-free data collections or publication of source code in order to supervise the market are not always possible and may result in so-called freezing effect. It seems justified to introduce such solutions for high-risk implementations but in our opinion using this solution for other type of implementations is not justified. We hereby express our concern whether incorrect interpretation resulting from incomprehensive or non-representational database has the same negative result as i.e. interpretation of medical testing.

We find the term “safety component” from para. 3 section 14 unclear, especially in connection with the radio devices directive. Therefore, we find it justified to clarify whether, for example, Android system is a “safety component” of a mobile device and whether the obligation means that the device itself or its system must perform some function that is critical for security.

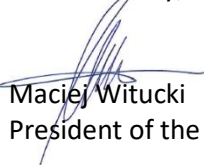
11. The creation of the European Artificial Intelligence Board

We advise including a balanced industry representation in the envisaged European Artificial Intelligence Board which has not only an advisory role but can have a real impact on decisions made, as well. It is important that this new body should become a platform of a true public-private collaboration and be composed of industry representatives on an equal footing with public stakeholders.

12. The post-market monitoring of an AI system

The Art. 61 concerning the post-market monitoring of an AI system should be removed or significantly reduced since it implies only additional burden for providers. Additionally, not all AI systems could be monitored (e.g. a product with an integrated AI system like an autonomous car).

Your faithfully,



Maciej Witucki
President of the Polish Confederation Lewiatan

