

Preliminary position of the Slovak Republic on the “White Paper on Artificial Intelligence – A European approach to excellence and trust”

Executive Summary

The Slovak Republic welcomes and supports the White Paper on Artificial Intelligence - a European approach to excellence and trust. We are ready to cooperate on building solid ground for the future uptake of AI in a trustworthy manner. Our position is presented in the form of preliminary views concerning the two ecosystems of AI proposed by the European Commission.

Ecosystem of excellence

- We should find the right incentives to motivate the top AI researchers to conduct their research in the EU and to contribute to the growth of European research centres.
- In order to achieve excellence, European research should focus on selected sectors including the research in Natural Language Processing.
- We support the Digital Education Action Plan 2021-2027 to increase awareness of AI at all levels of education, including lifelong learning programmes.
- We place a strong focus on the issue of women's under-representation in the tech sector, including the domain of AI.
- We recognise the strategic importance of all initiatives focused on achieving a sovereign European AI infrastructure.
- We should empower SMEs developing AI as well as non-technological SMEs by the network of European Digital Innovation Hubs (EDIHs) and by services of dedicated trust officers.
- We need to build upon our existing partnerships with all companies irrespective of their size or revenue, including IT corporations and large tech companies.

Ecosystem of trust

- We support the idea that Ethics Guidelines for Trustworthy AI enriched by fundamental rights impact assessment should play a pivotal role in guiding the ethical design of AI applications.
- To classify AI applications as high-risk we should take into consideration the severity and probability of occurrence of significant harms disregarding the sector of AI application.
- We should motivate developers to build explainable and transparent AI, especially for high-risk AI applications. During the development and production lifecycle there should exist a responsible role accountable for the explanation of AI decisions.
- We believe that backward-looking types of responsibility should be accompanied by the *ex-ante* dimensions. However, we still need to elaborate on a precise set of rules for taking responsibility.
- We support the hybrid approach to regulation of AI - direct regulation of high-risk AI applications and common self-regulation practices (accompanied with shared European labelling scheme) for low-risk AI applications.

Comments of the Slovak Republic on the White Paper on Artificial Intelligence – A European approach to excellence and trust

1. Introduction

The Slovak Republic welcomes the aim of the White Paper on Artificial Intelligence - a European approach to excellence and trust (doc. COM (2020) 65 final) to jointly build and execute a coordinated European approach towards achieving human-centric and trustworthy AI. We are ready to cooperate on establishing a regulatory and investment oriented approach aiming to promote the European research, development and uptake of trustworthy AI by businesses, public sector institutions, individuals and all relevant societies and organizations while tackling the risks associated with certain applications of this technology.

Below, we present our preliminary views on the main issues put forward in the White Paper in order to contribute to the debate on policy and legislative options that would enable a trustworthy and secure development of AI in Europe, with full respect of the values and rights of EU citizens. We welcome and support the division of this complex topic into two main building blocks - the “ecosystem of excellence” and the “ecosystem of trust”. We will provide more detailed views and positions reflecting the discussions on the national level on AI aspects raised in the White Paper and beyond upcoming discussions with the European Commission and with the other EU Member States.

2. An ecosystem of excellence

The Slovak Republic agrees with the view that Europe needs a more efficient approach towards building a strong position in AI research. We should **motivate the top European researchers in the field to stay in Europe** and create the most accommodating environment for their work. In addition, we should **motivate prominent non-EU researchers to join European research teams and to foster our AI research**.

Furthermore, we need to strengthen the position of European research centres. Currently, Europe has only three centres in the global TOP 25 research AI centres [1]. We **should focus on at least doubling the number of elite AI research centres** as well as to increase the number of top AI researchers in Europe in the forthcoming years. In coherence with the idea of building trustworthy AI applications, we **should cultivate the culture of ethical research and development in these centres**.

The Slovak Republic supports the approach that European AI research should focus on selected sectors (e.g. healthcare, transport, finance, agrifood value chains). We would like to emphasize that due to the linguistic diversity in Europe, investing into domains like **Natural Language Processing (NLP) would also help to initiate further research** in affiliated AI topics and to spark the interest of the general public.

The 21st century is a century run on talent. We already suffer from competence shortages across the globe. We therefore **support the reinforcement of the Skills Agenda, and welcome the Digital Education Action Plan 2021-2027**. Defining digital literacy as essential for life in a digitalised world allows us to approach the paradigm of how to keep educating while transforming with the right bias towards acceleration of development of digital curriculums, digital delivery of its contents, digital methodologies and virtual/augmented reality educational tools. We agree with the identified distinction between basic digital skills and the latest advanced digital skills in the Digital Education Action Plan 2021-2027. Given the size of the population of the Slovak Republic, we are well positioned to take full advantage of the proposed agile development of AI related policy and practice. We are committed and need to use our best efforts to **increase awareness of AI at all levels of education, including lifelong learning programmes**, in order to prepare citizens to make informed decisions and to actively participate in the AI era.

In the Slovak Republic, we have a strong focus on the issue of **women's under representation in the tech sector**, where women hold only 13.7% of the sector jobs. Women represent a majority of our population, and in order to provide for a pipeline of talent robust enough to address the acute global talent shortage and facilitate activities such as transformation of the assessment list of the ethical guidelines into an indicative "curriculum" for developers of AI, we think, that **solving the issue of getting more girls and women interested in ICT subjects and careers is the way forward**.

The AI ecosystem has a physical backbone. A **full stack European infrastructure of state of the art secure and sovereign data centres, data routes and data transfer protocols** are at the heart of our ability to ensure the functioning of the common European data space. The trend of exponential data growth with an estimated 175 ZB of stored data globally by 2025 challenges us to seek energy efficient and high-quality data storage and processing solutions. We welcome initiatives underpinning the European data spaces, such as GreenData4All and Destination Earth.

The Slovak Republic supports the idea of building collaboration networks of SMEs via specialized AI Digital Innovation Hubs (DIHs). The support should **empower not only the community of AI developers, but also non-technological SMEs prepared to boost their performance** with AI products and services. The Slovak Republic has proposed a manual aiming to help SMEs to embrace AI-based applications in selected segments [2]. Moreover, we need to encourage close cooperation between the AI DIHs in the EU. The Slovak Republic would welcome if a **communication platform spreading the know-how across the whole European DIH network** was developed and facilitated by the European Commission.

The Slovak Republic agrees that AI applications should be trustworthy for users. In view of this, **we need to educate and support the developers from SMEs to incorporate the ideas of trustworthy AI** into their products and services. Therefore, an inherent part of the AI DIH services should be a **service from dedicated trust officers** who will support the ideas of ethical AI design from early R&D stages.

Yet, most of the advanced AI innovation is pursued by IT corporations and large tech oriented companies. That is why we should **build upon our existing partnerships with all companies irrespective of their size or revenue**. We should not exclude anyone of good will who is ready to share knowledge. Good cooperation between technological leaders, IT corporations, SMEs, public sector and academia and sharing of know-how and datasets would help further develop AI in the EU. Balanced public-private partnerships can significantly help to establish better conditions for the development of trustworthy AI systems, especially for the use in the public sector but also for the deployment and use on a broader scale e.g. in the social media sector.

3. Ecosystem of trust

The Slovak Republic subscribes to the view that Ethics Guidelines for Trustworthy AI (Ethics Guidelines) should play a pivotal role in assessing the ethical aspects of AI applications. The Ethics Guidelines have the potential to be the first source of information and the reference point for AI developers and deployers regarding the ethical design of AI driven services and products. The Slovak Republic will welcome all initiatives that will help to promote the value of the Ethics Guidelines and improve its implementation. We have recently initiated the **establishment of the national AI ethics committee**. One of its roles will be the transformation of the Ethics Guidelines into an indicative “curriculum” for AI developers in The Slovak Republic.

We understand that AI can do much good but it can also cause harm. Whereas the material harms can be easier to grasp and understand, the immaterial harms, such as infringements of fundamental rights, need to be stated more clearly. The Assessment List for trustworthy AI (ALTAI)[3] which stems from the Ethics Guidelines, states that prior to self-assessing an AI application a fundamental rights impact assessment (FRIA) should be performed. The Slovak Republic therefore proposes to **establish a clear and unified process of FRIA for AI actors and stakeholders as an inherent part of ALTAI**.

All AI applications should be assessed with appropriate levels of human oversight. Especially **in case of high-risk applications there should always be a human capability to control the decisions made by AI algorithms**. The Slovak Republic fully supports the culture of human-centric, responsible and sustainable innovations. In general, in case of high-risk AI applications the more requirements of trustworthy AI we can reasonably impose and fulfil the more beneficial and less harmful such AI solutions can be to the society, especially when used on a broader scale affecting numerous people and for public interest.

The Slovak Republic states that the approach from the White Paper on AI where AI high-risk applications are linked with predefined “high-risk sectors” (e.g. healthcare, transport or energy sector) would be difficult to capture, since even implementations from the “low-risk sectors” can be potentially very dangerous for stakeholders. We would welcome a clear use-case based delineation between high-risk AI applications and those that do not pose such a risk to stakeholders. **To classify certain AI applications as high-risk we should take into consideration also the severity and probability of occurrence of such negative impact.** The Slovak Republic is convinced that a clear definition of high-risk AI applications

accompanied with a rich and open catalogue of specific use-cases will be beneficial to all actors and stakeholders.

We admit that it can be very difficult to predict the number of high-risk AI applications and it is possible that at the beginning there could be a lot of them. Therefore, the Slovak Republic proposes to focus on helping AI developers to better understand the ethical and regulatory boundaries within which they can develop their applications. **The Slovak Republic would welcome further development of an interactive European trustworthy AI risk-assessment tool** as a part of ALTAI, which would help the AI developers, providers and public bodies to evaluate and prioritize ethical risks with more precision (e.g. inspired by already existing frameworks [4]).

The Slovak Republic welcomes initiatives for transparent AI, especially in the field of high-risk AI applications. We need to **motivate developers and deployers to notify stakeholders if specific decisions were made not by human actors but (fully or partially) by AI algorithms** and if there is an alternative for a non-AI (human) solution.

The Slovak Republic also supports the view that AI models should be explainable at a reasonable level. We do not require that AI algorithms have to be understandable for every stakeholder at every stage of the development process. Yet, **during the production lifecycle there should exist a responsible role accountable for the explanation of AI decisions** that affect the stakeholders. Especially, in case of high-risk applications we should strive for the best possible explainability. The Slovak Republic **approves further research, development and application of explainable AI (XAI) in proper combination of ex-ante and ex-post principles.**

To a great extent, constructive societal trust in AI will be built on a track record of its verifiable beneficial societal outcomes. The ability of the public sector to help increase the degree of explainability of AI systems via policies, regulatory framework and practice will play a crucial role in creating and promoting individual beneficial societal outcomes over time. Therefore, we specifically support the view that **AI decisions for high-risk applications in the public sector should be explainable.** Moreover, we should motivate AI developers in the public sector to focus on **ex-ante** approaches to explainability and prefer the explainability by design for high-risk AI applications.

We agree with the notion that AI deployment in the public sector may have the transformative potential, as stated in the report by the Joint Research Centre (JRC) [5]. However, positive impact is far from straightforward and should not be taken for granted. Research by the JRC shows that the **deployment of black box solutions within the public sector carries a high risk of lacking legitimacy** and its use can be stopped by the judiciary or other law enforcement and control mechanisms. The rise of deployment of isolated AI Proof of Concept solutions can also lead to difficulties in materialising the return on investment when the solutions need to be maintained, scaled up and integrated to feed the open data datasets. **We need to support more testing and experimentation of AI solutions in governing public matters before we decide to scale them to real world practical application.**

We will need to clearly **define rules for taking backward-looking responsibility (fault-based or strict liability) for actions and decisions based on AI algorithms** where it is not easy to determine the offender or causal chain. We need to agree upon in what context and why we should broaden the scope of strict liability put on the developers and deployers and how to define the accountability roles in the production cycle [6]. We believe that the scope of strict liability in AI should be based on expert discussion, broad agreement of stakeholder groups and based on empirical data from affiliated use-cases (e.g. motor vehicles or aircrafts). We by no means want to hamper innovation but **we need to be sure that clear rules for responsibility ascription will be available** and applied fairly for all actors.

The Slovak Republic believes that **backward-looking types of responsibility should be accompanied also by the ex-ante dimensions** based on clearly stated obligations and the ethical virtues of AI teams [7]. Clear definition of forward-looking obligations and building the culture of responsibility alongside development and production chains would help to proactively mitigate most of the negative impacts of AI and set trustworthy environments for end-users as well as transparent conditions for AI developers and deployers.

Concerning regulatory framework, the Slovak Republic supports **the hybrid approach combining regulation and self-regulation of AI**. That means **direct regulation of high-risk AI applications and in case of AI applications which will not be assessed as high-risk to apply common self-regulation practices**. To rely exclusively on self-regulation could in some cases lead to irreversible damages and harms. We must **make sure that the relevant regulations have been met by all high-risk AI applications** before entering the market.

In view of this, we would welcome the creation of **common European certification procedures for trustworthy AI**. The certification should be iterative aiming to motivate all AI developers and deployers, including SMEs, to join the program for trustworthy AI and acquire proper sensitivity in the development of trustworthy AI applications.

The Slovak Republic **supports the idea of voluntary European labelling schemes**. These schemes would motivate the companies developing low-risk AI applications to aim for trustworthy and ethical AI. With the Seal of Trustworthiness they could acquire significant competitive advantage in the market.

[1] <https://macropolo.org/digital-projects/the-global-ai-talent-tracker/>

[2] <https://www.mirri.gov.sk/wp-content/uploads/2020/03/Dielo2-Manual.pdf>

[3] https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=68342

[4] <https://ethicstoolkit.ai/>

[5]

https://publications.jrc.ec.europa.eu/repository/bitstream/JRC120399/jrc120399_misuraca-ai-watc_public-services_30062020_def.pdf

[6]

<https://ec.europa.eu/info/publications/commission-report-safety-and-liability-implications-ai-interne> and
https://www.europarl.europa.eu/doceo/document/JURI-PR-650556_EN.pdf

[7]

https://ec.europa.eu/info/news/new-recommendations-for-a-safe-and-ethical-transition-towards-driverless-mobility-2020-sep-18_en