Daan van Pinxteren - 15 October 2024*

The EU AI Act: The Impact on the Public Sector

Introduction

On 2 August 2024, after a lengthy negotiation process, the EU Artificial Intelligence Act (AI Act), the world's first comprehensive regulation on AI came into force. With this new law, the EU aims to protect fundamental rights by facilitating the development of trustworthy AI, while at the same time encouraging innovation with the technology.[1] The public sector, which in recent times increasingly depends on AI to enhance its services, is largely impacted by the Act. This leads to several questions: how will the AI Act restrict the public sector use of AI to provide more efficient, effective and tailored services to citizens? How will it facilitate the development of trustworthy AI, which is key for the public sector, as the unregulated use of AI may lead to unfair and discriminatory outcomes with dire consequences for society?[2] In turn: to what extent will this lead to compliance with the new rules? To shed light on these questions, we must first consider two main characteristics of the AI Act.

1) A risk-based framework

The distinctive feature of the new AI Act is that it follows a risk-based approach related to the risk AI systems pose to the fundamental rights and freedoms of individuals. In this regard, the AI Act sets out different risk levels, starting from unacceptable risk systems that are entirely prohibited to limited risk systems such as chatbots and minimal risk systems such as video games and spam filters. Particular attention is also given to general purpose AI (GPAI) models (including generative AI), which now form a separate category under the AI Act after their rapid rise in recent years. Most rules are, however, directed at high-risk AI systems, which are systems that the EU deems risky either because they are (safety components of) products that fall under EU harmonisation legislation, such as medical devices and toys, or because they are part of a list managed by the Commission that sets out systems that are likely to impact fundamental rights and freedoms of citizens.[3] Various use cases of public sector AI will fall under this latter category and for these systems, there will be many new obligations for organisations starting from 2 August 2026.[4]

2) Obligations across the value chain

The kind of measures an organisation needs to take is dependent not only on the risk category, but also on the role that it has in the AI value chain. Most requirements are directed towards 'providers', which develop AI systems or place them in service under their own name. For limited risk systems and GPAI models, obligations for these parties are mostly focused on keeping documentation and providing transparency. Yet, for high-risk systems, there are many additional requirements, such as the implementation of risk management and quality management systems to estimate and evaluate risks when the AI system is used, drafting technical documentation to demonstrate compliance, and conformity assessments to ensure systems adhere to the rules.[5] Public authorities are deemed 'providers' of AI systems if they develop their own AI systems or purchase tailor-made systems.

When public authorities actually use AI systems (whether they develop them themselves or not) they are considered to be a 'deployer' under the Act, which means there are additional requirements. For high-risk systems, the deployer i.a. needs to implement technical and organisational security measures and assign mandatory human oversight with the necessary competence, training and authority.[6] Public sector deployers also need to conduct a Fundamental Rights Impact Assessment (FRIA) to evaluate risks to individuals and mitigate these risks[6] and register the use of AI high-risk AI systems in an EU database.[8]

* **Daan van Pinxteren** *is working as a consultant in a large advisory firm. In this role, he gives advice to the public and private sector on digital laws and regulations, with a focus on privacy, data ethics and AI. Daan has recently graduated from CIFE's Executive Master in EU Studies where he did his final research on the AI Act and its impacts on the Public Sector. Daan also has an academic background in law and economics.*

The dichotomy of risks in the AI Act

One key factor that can be extracted from these core characteristics of the AI Act is that there exists a clear dichotomy in its risk classification system: stringent rules for high-risk systems and the prohibition of unacceptable risk systems, but only minor requirements, i.e. transparency requirements for other AI systems.[9] For some Public sector AI systems, such as those performing limited administrative tasks, this is good news, as these only require transparency. Yet, for the wide range of applications of AI in the Public sector that fall under the high-risk category of AI, public authorities that provide or deploy AI have many responsibilities, potentially leading to high costs of compliance or even prohibitions.

This dichotomy becomes clear by investigating use cases of AI in the area of public administration. In this context, AI is particularly useful for automating routine administrative tasks, freeing up civil servants to focus on more complex and high-value activities. For these systems, the AI Act will have a minimal impact as these will most likely fall under the limited risk category. However, more sophisticated AI systems that perform complex tasks, such as models involving profiling to determine if citizens are eligible for certain services, such as financial aid, will likely fall under the high-risk category, and then a variety of measures would need to be implemented.[10]

Thus, the increased efficiency and effectiveness that AI systems in the public sector may bring may be offset by the high cost of compliance with the AI Act in certain situations. The question then arises: will the high costs of compliance be counterbalanced by other benefits of the rules, such as the facilitation of the development of trustworthy AI and its contribution to citizens' trust?

Trustworthy AI and citizen trust

The AI Act started out as an instrument to protect the fundamental rights of individuals that are threatened by AI.[11] However, due to the many compromises during negotiations and pressure to preserve innovation, this focus on the protection of fundamental rights and the development of trustworthy AI has diminished in the final text. The final text mostly focuses on market access and product safety and some consider the human rights aspect to be more of an afterthought.[12] Thus, although a key goal of the EU was to facilitate the development of trustworthy AI, several aspects of the rules detract from this goal. Furthermore, current research shows that trustworthiness of AI systems may not actually lead to more trust in governments. First of all because the actual effects of transparent and accountable AI on trust are unclear as this topic has not been sufficiently researched in the EU. Second, trustworthiness is not simply a matter of risk-grading, which is currently the central focus of the AI Act. Finally, trust has an irrational component as it is not based on purely rational deliberations, for example, the fear of what AI is capable of may play a role in trusting it, even if it is, in fact, more trustworthy than a human.[13]

With this, the benefits of the AI Act for the Public sector in facilitating trustworthy AI and in turn increasing public trust may be limited, again compared to the actual costs of compliance. This imbalance is further complicated by the expected enforcement of the rules.

The effectiveness of monitoring and enforcement

Under the AI Act, organisations themselves classify the risk of their AI systems. Then, after determining the risk category, it is again up to organisations to ex-ante determine if their systems comply with the measures described in the rules and subsequently apply these measures. Hence, the assessments that need to be made, such as the conformity assessment or a FRIA, are self-assessments without external control. Moreover, the type of risk management system to be set up, or the security measures to be implemented, are up to the discretion of the organisation, and may thus differ in quality.[14]

Moving towards external enforcement, the supranational AI Office and AI Board, which oversee the correct application of the rules across the EU, do not have any enforcement powers. Member States themselves need to designate authorities responsible for monitoring organisations and the enforcement of the rules, i.e. notifying bodies and market surveillance authorities.[15] This Member State discretion may lead to the uneven implementation of enforcement across the EU and it is as yet unclear how these enforcement bodies will work in practice, whether they will be efficient and whether their powers will be efficient.[16] To go even further, the AI Act specifies that Member States themselves should lay down rules regarding the fines imposed on organisations. However, for

public authorities, Member States can lower these or abolish them altogether.[17]

Lastly, as mentioned before, the focus of the AI Act lies on conformity and risk rather than the protection of individuals. This is also shown in the enforcement mechanisms; individuals have no right to redress or to raise a complaint under the rules and only those with obligations under the AI Act can challenge regulators' decisions, not those whose rights are impacted.[18]

Is non-compliance an option?

Owing to the anticipated high costs of compliance for high-risk systems, together with the possible limited positive effects on citizen trust and the expected scattered enforcement of the rules towards public authorities, it may be tempting for these bodies to move towards non-compliance with the rules (for example by purposely misclassifying their AI systems as limited risk, so avoiding the stringent measures). In this regard, a political factor may also come into play, since disruptions in efficiency and effectiveness with AI due to the burden of compliance may open pathways to entirely different scenarios, namely a perception that when current (democratic) systems are not efficient, the allure of simplified, yet undemocratic solutions may gain traction - such as those purported by populist and authoritarian regimes that disregard or oppose rules on AI.[19]

Yet, with the previously outlined risks of unregulated AI in the public sector, the importance of public values in public services, such as transparency, privacy, non-discrimination and inclusivity and the exemplary role that public authorities have regarding this, non-compliance with the Act may not be a valid option. Rather, public authorities may consider shifting their AI strategies from developing and implementing high-risk applications towards AI use cases that have only limited risk, as positive effects on effectiveness and efficiency are then less outweighed by the high costs of compliance resulting from the many measures that need to be taken. However, some critics argue that although even when? systems are deemed non-high-risk under the AI Act, they may still carry risks to the health, safety and fundamental rights of individuals.[20] Furthermore, those systems, performing only administrative tasks, may lead to less efficiency and effectiveness than (more advanced) systems that fall under the high-risk category. Luckily for public authorities, the AI Act also offers some leeway in this regard.

Harnessing exceptions to the rules

In some specific AI use cases in the public sector, there are exceptions that authorities can harness. One such exception occurs in the context of law enforcement.[21] For example, the human oversight requirement is somewhat less stringent for law enforcement agencies, with no mandatory verification of individuals in the context of remote biometric identification.[22] In addition, although some systems in the context of law enforcement are completely prohibited, for example those related to real-time remote biometric identification systems in publicly accessible spaces, there are again some exemptions, as systems specifically used for a targeted search of a missing person or preventing a terrorist attack are exempted from the prohibitions.[23]

Something similar can be said for AI use cases in the area of national security and defence as the nature of national security as a responsibility of Member States means that these areas are excluded of the scope of the AI Act.[24] Yet, Public authorities should be aware that the national security exemptions do not always apply. This is because AI systems in these contexts are often dual-use, with both civilian and military purposes (e.g. drones).[25] In these dual-use cases, the AI Act would apply to the systems (as it is not solely used for national security and defence) and may be categorised as high risk, with many requirements.

In addition to the exemptions outlined above, there are quite a number of other exceptions in the AI Act which may offer leeway to public authorities. Therefore, these institutions would do well to familiarise themselves with these to optimally navigate the rules.

Conclusion

The EU Artificial Intelligence Act will significantly influence the public sector's use of AI. While low-risk systems may continue to enhance efficiency and effectiveness with minimal regulatory impact, high-risk systems face stringent requirements that burden public authorities with compliance challenges when these come into force in August 2026. These compliance challenges are further underlined by the possible

lack of benefits of compliance, i.e. fostering trustworthy AI and citizens' trust and the possible scattered and ineffective enforcement of the regulation, reducing the incentive to comply.

However, non-compliance with the rules to harness effectiveness and efficiency with AI is not really a feasible option for the public sector. Instead, public authorities should consider reinventing their AI strategies, striking a balance between effectiveness and efficiency and minimal costs of compliance, leveraging the difficult concepts of trustworthy AI and citizens' trust, or harnessing the many exceptions to the rules. From the perspective of the law itself, the success of the AI Act and compliance with it by the public sector will ultimately thus depend on its ability to foster efficiency and effectiveness with AI without a too higher burden of compliance, and its ability to actually safeguard democratic values in the development of AI, ensuring that AI serves the public good rather than undermining it.

References

1 See European Commission, "European Artificial Intelligence Act comes into force", European Commission Press Release, 1 August 2024.
2 The Dutch tax fraud algorithm scandal is exemplary to this as it has led to much criticism regarding the use of AI by governments, since it may violate principles such as non-discrimination. See Heikkilä Melissa, "Dutch scandal serves as a warning from Europe over risks of using algorithms", Politico, 29 March 2022.
3 Article 6 and Annex I and III of the AI Act.
4 Obligations do not apply for high-risk systems put on the market before this date, except for systems provided or deployed by public authorities, that need to comply from 2 August 2030 onwards. See Article 111 of the AI Act.
5 Articles 8 to 18 and article 43 of the AI Act.
6 Articles 26 of the AI Act.
7 Similar to Data Protection Impact Assessments under the GDPR. See article 27 of the AI Act.
8 Article 49.3 of the AI Act.
9 Wörsdörfer, Manuel, "Mitigating the adverse effects of AI with the European Union's artificial intelligence act: Hype or hope?", forthcoming in: Global Business and Organisational Excellence, 43(3) (2024), p. 22.
10 Annex III 5(a) of the Draft AI Act.
11 See for example European Commission Press Release, 7 December 2018 European Commission, "Europe fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence", European Commission Press Release, 21 April 2021.
12 See note 8 above, p. 19.
13 Laux, Johann, Sandra Wachter and Brent Mittelstadt, "Trustworthy artificial intelligence and the European Union AI act: On the conflation of trustworthiness and acceptability of risk", Regulation & Governance 18, no. 1 (2024), pp. 4, 6 and 24-26.
14 See Smuha, Nathalie A., et al., "How the EU can achieve legally trustworthy AI: a response to the European Commission's proposal for an Artificial Intelligence Act," LEADS Lab @University of Birmingham (2021), p. 37-39.
15 See for example Articles 31 and 70 of the AI Act.
16 See note 13 above, p. 46-48.
17 Article 99 of the AI Act.
18 See note 13 above, p. 44-46.
19 Peixoto, Tiago C., Otaviano Canuto, Luke Jordan, "AI and the future of government: unexpected effects and critical challenges", Policy Center for the New South, 20 March 2024.
20 De Cooman, Jerome, "Humpty dumpty and high-risk AI systems: the ratione materiae dimension of the proposal for an EU artificial intelligence act." Mkt. & Competition L. Rev. 6 (2022), pp. 63-64.
21 due to pressure of the council during negotiations to obtain leeway for law enforcement agencies, being a strong Member State competence.
22 Article 14.5 of the AI Act.
23 Article 5.1(h) of the AI Act. European Parliament, "Artificial Intelligence Act: MEPs adopt landmark law", European Parliament Press Release, 13 March 2024.
24 Recital 24 of the AI Act argues that this exclusion is justified by Article 4(2) TEU and by the common EU defence policy which are subject to public international law and is therefore the more appropriate framework to regulate these type of AI uses.
25 See for example. Carrozza, Ilaria, Nicholas Marsh and Gregory M. Reichberg, Dual-Use AI Technology in China, the US and the EU, PRIO Paper 2022.

Rédaction: Laurent Baechler et Aline Palige
Policy Paper / Note de recherche est publiée par le Centre international de formation européenne, association dont le siège est 81, rue de France, F-06000-Nice.

www.cife.eu

Ce projet a été financé avec le soutien de la Commission européenne. Cette publication (communication) n'engage que son auteur et la Commission n'est pas responsable de l'usage qui pourrait être fait des informations qui y sont contenues.

Avec le soutien du programme Erasmus+