Towards Trustworthy AI in the EU

Mastercard's position on the Artificial Intelligence Act

Mastercard welcomes the opportunity to provide feedback to the European Commission on the Proposal for a Regulation laying down harmonised rules on artificial intelligence (the "Artificial Intelligence Act" or "AI Act").

This consultation comes at a difficult moment for many individuals, communities, and businesses. Nearly a year and a half into the COVID-19 pandemic, despite breakthrough vaccines, we are still fighting the tragic health implications of the virus and grappling with the new reality of physically distanced and increasingly digital lives.

At a moment of increased vulnerability across the economy, AI-based tools, built on a robust governance framework, have helped Mastercard facilitate access to the digital economy, thus increasing trust in our payment network and solutions.

Mastercard understands the importance of trust and decency. Clearly the events of the past 18 months have only underscored the need for these values in our lives and business relationships. Mastercard therefore fully supports the creation of a proportionate regulatory framework that will facilitate the uptake of trustworthy AI in the European Union and wishes to respectfully submit its comments on the following main topics:

1. Definitions and scope
2. Remote biometric identification
3. Requirements for high-risk AI systems
4. Enforcement and competent authorities

1. Definitions and Scope

a. AI System

The AI Act defines an 'AI system' as "*software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with*".

It has been our experience that certain techniques and approaches listed in Annex I actually involve limited risk in practice, i.e., only negligible impact on the health, safety or fundamental rights of persons. Indeed, certain standard statistical or logic-based techniques, e.g., regression analysis or t-test analysis, have been around for years without engendering risks that cannot be addressed by the existing regulatory framework. In fact, certain such approaches deliver considerable value, such as a logic-based approach used to send alerts to cardholders that their next transaction may not go through because their card is approaching the credit limit, with the aim of protecting them from going into debt. Similarly, an AI technique used to provide suggestions to cardholders when they are setting spending controls for their card (e.g., I do not want transactions of 100+ EUR between 11pm and 5am to be approved) based on insights into spending patterns and the how and when of fraudulent transactions protects cardholders without significant harmful impact on their health, safety or fundamental rights.

We understand the Commission's regulatory intent to target techniques and approaches that could lead to new or incremental risks stemming from AI, including a certain level of opacity, whereby an input given to an AI system leads to an output that is not (fully) understandable.

However, in order to ensure consistent application, the "output generation" element of the definition should be clarified in such a way that makes it apparent that it covers **system-generated AI outcomes**; i.e., AI systems whose outputs are generated based on rules stemming from the AI system itself, not human-generated ones, such as a sophisticated excel sheet the logic of which is fully developed and controlled by humans.

b. High-Risk AI

Mastercard welcomes the AI Act's risk-based approach and its focus on high-risk AI use cases listed in Annex III, which is needed to protect against harm and increase trust in AI systems.

Nevertheless, the AI Act should consider that the level of risk of an AI system depends on the specific circumstances of the AI use case. For instance:

- AI used for recruitment can be sensitive, but there may be recruitment use cases where the risk is low because there is no appreciable impact on future career prospects and livelihoods. Take the example of an AI system that scans whether the word "python" in a CV is used to refer to an individual's Python programming language skills rather than their work experience at a zoo with a particular type of snake, wherein the scan would take place to help identify relevant skills for a recruiter to evaluate.
- Banks using AI to monitor their total line of credit in accordance with legal requirements entails lower risk compared to AI used to evaluate creditworthiness when granting credit to an individual. In particular, when credit is already granted to a number of individuals or businesses and AI is used to understand, measure and manage risk related to a bank's overall credit portfolio in order to determine whether additional credit can be extended to the individual or business because they are doing well. Similarly, when a bank considers onboarding a new business, it can rely on AI to offset creditworthiness of said business against its overall credit portfolio and broader risk profile to help the bank make an informed decision without necessarily determining individuals' access to financial resources.

The list in Annex III should therefore be **further defined**, to ensure that those AI systems listed actually do pose a high risk and to avoid the unintended inclusion of low-risk systems. By applying the most burdensome requirements to only a select list of high-risk AI systems, the EU would achieve its aim of protecting individuals while also supporting AI-based innovation, especially when it comes to lower-risk AI applications.

2. Remote Biometric Identification

The AI Act qualifies as high-risk "*AI systems intended to be used for remote biometric identification of persons*", making such systems subject to the requirements of Title III, Chapter 2 of the AI Act.

We understand this aims to cover remote biometric *identification* as opposed to biometric *authentication*. This is consistent with the Commission's February 2020 AI White Paper (in footnote 52), which specified that "*remote biometric identification should be distinguished from biometric authentication (the latter is a security process that relies on the unique biological characteristics of an individual to verify that he/she is who he/she says he/she is)*."

The reason for this distinction is clear. Remote biometric identification can be particularly intrusive from a fundamental rights standpoint when it:

(i) involves the processing of biometrics of an **indiscriminate number** of individuals and requires comparing an individual's biometric data to the biometric data of many other individuals stored in a database to identify said individual (i.e., one-to-many matching). On the other hand, biometric authentication entails less risk, given that it consists of comparing two biometric templates usually assumed to belong to the same individual (i.e., one-to-one matching) and no link with the actual identity is established.

(ii) covers situations where biometric data is used to identify individuals **without their knowledge**, rather than a situation where a well-informed individual deliberately chooses to confirm their identity based on their biometrics in order to transact or otherwise interact with a service provider, such as through an e-commerce platform, in-person shop or an online government service.

The definition of 'remote biometric identification system' in the AI Act currently reads: "*an AI system for the purpose of identifying natural persons at a distance through the comparison of a person's biometric data with the biometric data contained in a reference database, and without prior knowledge of the user of the AI system whether the person will be present and can be identified*".

We recognise the Commission's intent to have a functional definition irrespective of the technology, processes or types of biometric data used. However, in our view, this does not make a clear enough distinction between remote biometric identification and biometric authentication, and risks unintentionally impacting identity verification systems relying on biometric authentication or other security procedures for which knowledge of the actual identity is not needed. The definition of remote biometric identification system should therefore be clarified, clearly excluding biometric authentication.

Furthermore, it would be worth specifying that "*without prior knowledge of the user of the AI system whether the person will be present and can be identified*" refers to scenarios where data of an indiscriminate number of individuals is processed - unbeknownst to those individuals - for the identification of only a few. It is important that the two conditions of (i) indiscriminate character, and (ii) without prior knowledge are clarified, and in that way be made distinct from a situation where individuals have received notice and actively opted into the use of a service (for example, in the digital economy, to file their tax declaration, move faster through an airport security queue, etc.).

3. Requirements for High-Risk AI Systems

The AI Act puts strict requirements in place for high-risk AI systems. While we agree that these requirements are important in certain AI contexts, we consider that some of the requirements are quite prescriptive and may not fully align with actual risks and trade-offs of an AI system.

We recognise that Article 8(2) of the AI Act allows organisations to take into account the intended purpose of the AI system and the risk management system when ensuring compliance with the AI Act's requirements applicable to high-risk AI systems, and that the risk management system itself will involve an evaluation of potential risks when the system is used "*in accordance with its intended purpose and under conditions of reasonably foreseeable misuse*" (Article 9(2)(b)). We understand this is meant to provide organisations the flexibility to adapt the requirements to the specific contexts in which the AI system can be deployed and the related risk level. Nevertheless, these provisions leave considerable room for interpretation by AI providers, users, competent authorities, and other actors in the AI value chain. In addition, notions such as "intended purpose" and "reasonably foreseeable misuse" often depend on actions by the user of the AI system rather than the provider.[1]

For example, the AI Act requires high-risk AI systems to be designed and developed with an appropriate level of accuracy, robustness, and cybersecurity (Article 15). Accuracy is an important principle to consider in AI applications. Yet, aiming for maximum accuracy of an AI system does not always improve the effectiveness of the AI output or help to protect against individual harm. In the financial sector, for instance, reducing accuracy on purpose to increase false positives, and thus flag more payment transactions as potentially risky, allows for extra due diligence steps in assessing fraud, which in turn protects individuals. Whereas, in the abstract, reduced accuracy would be perceived as problematic or risky, in certain circumstances it can be a conscious trade-off to achieve an important individual and societal benefit – in our case preventing fraud.

Another example is that of transparency. The AI Act requires providers of high-risk AI systems to draw up electronic instructions for use for the AI system's user (Article 13). Such instructions shall include characteristics, capabilities, and limitations of the AI system. Those instructions for use must also be published in the public EU database to be managed by the Commission (Article 51 and Annex VIII). Transparency as a cornerstone in trustworthy AI must be balanced with the need to ensure that intellectual property and proprietary information remain protected, and that malicious actors are not encouraged to bypass the AI system.[2] For instance, a user of a high-risk AI system used to address credit risk will benefit from a detailed technical description of how the model was developed and how it generates outcomes, so they can take informed decisions when using the AI system. However, it is not appropriate to share such information broadly with individuals, given it is overly technical and could reveal confidential information that malicious actors could use to circumvent the process.

The AI Act should unequivocally confirm that organisations have the **flexibility** to "calibrate" requirements according to the **specific context**, in particular the given AI use case and related risk evaluation, to ensure that the AI Act can achieve its dual purpose of protecting against harm and fostering AI-based innovation. This would lead to the following benefits: (i) future-proof and technology-neutral rules for AI; (ii) a level playing field for all market players to compete and innovate; and (iii) consistency with other laws and regulations, including the General Data Protection Regulation (GDPR).

4. Enforcement and Competent Authorities

As with all legislation, the AI Act's success will be dependent on the efficiency of its regulatory oversight and enforcement.

[1] See also Section 5(b) below.

[2] Article 70 of the AI Act does not provide sufficient reassurance in this respect, given that it only requires national competent authorities and notified bodies to respect confidentiality of information obtained in carrying out their tasks under the Act.

We understand that several authorities may have competences under the AI Act, amongst which national competent authorities, market surveillance authorities, authorities responsible for the supervision and enforcement of financial services legislation, etc.

In addition to the competences bestowed by the AI Act, other authorities may have competences depending on the specific circumstances of the AI use case, such as the data processed or the sector in which it is deployed. For instance, the data protection authorities ("DPAs") will be competent for AI applications involving the processing of personal data, including where such applications have a potential impact on individuals' rights and freedoms, in accordance with Article 1(2) of the GDPR.

Such a complex enforcement framework with many different authorities may lead to overlapping competences amongst authorities. This risks inconsistent interpretation and application of the AI Act as well as ineffective enforcement, which would undermine EU citizens' trust in AI and the digital economy.

a. Lead Authority

To ensure smooth regulatory oversight and to reduce unnecessary burden on the development of future trustworthy AI applications, we believe it is critical to have **one authority taking the lead and operating as a single point of contact** towards the organisation concerned. For purposes of legal certainty, the DPAs should retain general competence over AI applications involving the processing of personal data and/or impacting individuals' privacy and other fundamental rights. As such, they should be in a leading position and the single point of contact for the organisation in question, but with the possibility to collaborate with other authorities where necessary.

For instances where AI does not involve the processing of personal data (e.g., AI used for manufacturing processes), it could be envisaged that the authority with the most relevant expertise (e.g., because of the sector in which the AI is deployed) takes the lead and cooperates with other authorities where needed.

b. Consistent Application

The resources required to ensure effective cooperation between different competent authorities should not be underestimated. Take for instance Article 63 of the GDPR, which puts forward the consistency mechanism, according to which the DPAs must cooperate with each other and, where relevant, with the Commission to contribute to the consistent application of the GDPR. However, since the GDPR's entry into force, it has proven difficult and burdensome for the different DPAs to put this consistency mechanism into practice and collaborate in an efficient way. This has led to diverging opinions and interpretations, which runs counter to the GDPR's initial objective of harmonisation.

The AI Act's regulatory oversight could learn from this experience and build further on the GDPR's achievements. The AI Act should therefore clearly spell out how **cooperation amongst different authorities** at different levels will work in practice.

5. Other suggestions for clarification

Beyond the topics discussed above, Mastercard would like to put forward the following suggestions for clarification:

a. It should be clarified that the qualification of high-risk only applies to the specific use cases listed at the sub-level in Annex III of the AI Act and not to the entirety of the areas listed in said Annex.

b. The AI Act imposes most requirements on the provider of an AI system. This does not reflect the fact that many requirements can in practice only be managed by the user as the entity controlling the AI system's use and the specific context in which it is deployed. For instance, a provider cannot reasonably be expected

to anticipate all potential uses or what data the user will feed into the AI system to adequately comply with the risk management system (Article 9), data governance (Article 10) and human oversight (Article 14(4) and (5)) requirements. Similarly, a provider may not always have information about the extent to which the AI system will play a role in the user's decision-making processes or what potential safeguards the user may put in place. We believe a more suitable balance should be struck between the responsibilities and roles of providers and users, to ensure responsibilities are allocated to the entities best placed to determine the actual level of risk and appropriate safeguards for the AI system.

c. It is unclear when AI systems intended to be used for remote biometric identification should go through conformity assessment by a notified body vs. conformity assessment based on internal control under Article 43(1) of the AI Act. It appears providers who have applied harmonised standards or common specifications would be allowed to choose between either of the two procedures, but this should be further clarified to ensure legal certainty.

d. It is unclear whether the requirement for high-risk AI systems to bear a CE marking under Articles 19 and 49 of the AI Act applies both to AI systems that have gone through conformity assessment by a notified body and those that have gone through conformity assessment based on internal control. Additional clarification would help AI providers to better understand regulatory obligations.

e. Apart from consistency with the existing EU legal framework, consistency should be ensured between the AI Act and the various upcoming legislative measures at EU level, in particular on data sharing (e.g., Data Act, Data Governance Act, etc.) and liability for AI.

Mastercard would welcome an opportunity to discuss the above with the Commission at your earliest convenience and looks forward to engaging with the co-legislators as the proposal goes through the legislative process.