EDiMA

EDiMA position paper on Artificial Intelligence

EDiMA's members are fully committed to maximising the benefits of AI for Europe, while addressing its potential risks. As such, EDiMA supports the Commission's goals of promoting investment and innovation in trustworthy AI. AI-powered analysis offers the European Union a range of opportunities in a broad variety of economic sectors and society in general, and it should be at the centre of the EU digital agenda.

An ecosystem of trust

While the current and potential benefits of AI are numerous, concerns about AI are important and legitimate and should not be discounted. Developments in AI are advancing quickly and will have a transformative impact on our societies, so it is important to foster an ecosystem of trust.

That said, AI is not a new concept - it has been used for a very long time and has only garnered public attention recently as its usage grows in scale. AI and Machine Learning are software tools to help classify information, make recommendations and support optimisation in ways that enhance organisational efficiency and human decision-making. Some AI applications are only linked to industrial processes and do not affect individuals at all. For example, AI can be used to optimise the electricity consumption in a factory, to predict when maintenance should be performed on some machines or to optimise the manufacturing process. It is also important to note that the vast majority of AI applications do not carry high risks for human beings, and that risks vary based on multiple parameters. Obviously, this brings to light that the definition of "high risk" as such needs to be clarified.

Concerns related to the functioning of AI are mostly factored into the development stages of the technology. Taking potential risks into account and having strong processes in place can help to minimise risks, which is essential to creating high quality products and services that consumers use and trust. An example of such a process is checking the quality of the datasets within the model used to train it, and by performing rigorous testing. It is, however, essential that these processes are flexible in order to adapt to widely different applications of AI.

Going forward, it is helpful to have auditing internal processes in place e.g. to audit the level of risk of using a particular data source as training data for various tasks, rather than always focusing on the quality of the datasets in itself. This is because the quality of datasets is not inherently high or low – it is either well- or ill-suited to train AI models for certain tasks, and datasets cannot be audited independently from the model being used to train it. For example, a dataset of images of human faces that contained 70% male faces may yield a high quality face detection model but a biased detection model.

Deployers of AI powered technology also have a role to play in ensuring continued safety, and to limit the adverse impact on fundamental rights. Indeed, these can be misused - deliberately or not - in ways that can raise safety risks or risks of discrimination for the end-users. While the developer should strive to minimize risks in the initial stages of development, and inform the deployer of the capabilities and limitations of a technology, the deployer must be responsible for the correct deployment of a specific solution.

A workable AI definition

As the European Commission made clear in its AI White Paper, when coming up with a workable Artificial Intelligence definition, one of the main challenges is to allow sufficient flexibility to "accommodate technical progress" while at the same time be "precise enough to provide the necessary legal certainty". Whereas a broad framing of AI would threaten to encompass all contemporary software systems, a definition that is too narrow would lead to an overly descriptive concept of AI that is likely to rapidly become outdated due to the rapid pace of digital innovation.

We also caution the Commission with defining AI as broadly as the open ended category of "automated decision making" as indicated in the Inception Impact Assessment on Ethical AI. This would go against the initial, thoughtful direction proposed in the AI whitepaper that proposes to focus on the risk-based, double-criterion for sectorial and application/use-based AI technologies. If AI were defined as "automated decision making" for the purpose of the future AI regulation, it would create disproportionate and unjustified regulatory obligations that would not only deter development and deployment of AI-based applications in Europe, but also automated systems that do not pose any risk nor harms.

EDiMA believes that a number of issues needs to be taken into account when considering AI definition and classification:

1. *Risk based approach*: It is important to assess both the *severity* and *likelihood* of risk posed by specific AI applications. By focusing not only on the severity that a certain AI application might cause, but also on the likelihood that something negative happens as a result of its use, we can better understand the risk attached to the use of a specific AI application. It is also important to take into account the benefits that certain AI applications can have, while keeping in mind the harm involved in not using a certain AI application (opportunity cost).
2. *Intended use vs Distorted use:* A distinction should also be made between intended use and distorted use of AI applications. If AI creates harm because it has been manipulated by users, then is it the AI at fault or rather the users in question? Ex: users manipulating search algorithms to increase the visibility of illegal content on a marketplace. This is also why too much transparency on AI might be dangerous as it might be giving the tools for ill-intentioned users to perform these manipulations.
3. *Context-awareness*: Rather than focusing on technical specificities of an AI application, it is important to rather focus on the context in which it is used. This will allow us to better understand who is impacted and in what circumstances. This is also a more *human-centred approach,* similar to that followed by GDPR when it comes to the treatment of automated decision-making.

Definition of high-risk AI and risk assessment

We agree with the Commission that any new regulatory framework on AI should be risk-based. Imposing new legal requirements beyond applications that are clearly defined as high-risk threatens to chill innovation and investment in the European market. It also bears its own risks in limiting certain AI technologies which are themselves known to reduce harm. Existing legislation should also be factored in in its capacity to resolve potential issues created by non high-risk AI applications e.g. GDPR.

The concept of cumulative criteria, whereby an AI application is identified as a high-risk application if both the sector *and* the specific intended use involve significant risk, is workable provided such indicators are clear and targeted. At the same time they should be sufficiently nuanced to take into consideration the large diversity of AI and emphasise the need for proportionality. By removing vague and open-ended clauses from the risk definition, such as "exceptional instances" and "immaterial damage", legal certainty will significantly improve and avoid potential regulatory overreach (e.g. ambiguous definition of immaterial damage could have adverse liability implications and lead to potential abuse).

Furthermore, in order to ensure legal certainty, the process to decide whether new sectors should be added to the list of 'high-risk sectors' must be robust and transparent, taking into consideration the views of relevant experts and with extensive consultation with all stakeholders. The decision to add a sector to the list should be followed by a reasonable transition period to ensure preparedness for AI developers.

When it comes to a risk assessment of an individual AI application, context and prevalence matter. Without knowing the purpose for which AI is used, who could be impacted, the scale of the risk or likelihood of harm, the relative importance of factors such as discriminatory outcomes or AI inaccuracies is difficult to assess.

As is the case with many other technologies, AI development is complex and unintended consequences can happen during the process. Even with datasets that are well-suited to be used to train AI models for certain tasks and strong internal processes, something unintended can happen depending on the type of AI technology used. There are processes such as audits, testing, monitoring, reviews, employee training etc. which can be put in place to mitigate that risk. Companies employ such effective design measures in order to deliver the most appealing and competitive products to their customers - thereby earning their trust.

Reflecting on whether there is a need for new compulsory requirements being limited to high-risk applications (where the possible harm caused by the AI system is particularly high), EDiMA believes a risk-based approach is important. We also agree that low-risk applications should not be subject to a regulatory burden. The concept of establishing cumulative criteria to easily identify high-risks applications is workable provided such indicators are clear and proportionate, and sufficiently nuanced to take into consideration the large diversity of AI applications.

Regulatory oversight over such practices should focus on applications that carry the highest risk for users, bearing in mind the possible negative consequences that could result in a broader set of applications being regulated, such as the hampering of innovation, greater administrative burdens, but also the potential opportunity costs (i.e. where limiting the use of AI could lead to potentially bigger risks than those that the AI system itself could generate). Sufficient flexibility should also be considered, taking into account the rapid development of the AI and machine learning capabilities and applications**.**

Ex-ante conformity assessments

One of the best ways to ensure that AI is trustworthy, secure and respectful of European values and rules is a combination of ex-ante risk self-assessment and ex-post enforcement for high-risk AI applications. EDiMA believes that ex-ante conformity assessment requirements alone do not strike the right balance, as they would be unlikely to achieve similar results in as fast manner without unduly halting innovation and creating unnecessary burdens.

The approach taken in GDPR can provide important lessons. By leaving behind the ex-ante approach followed by its predecessor, the GDPR put in place a stronger focus on the principle of accountability. The Data Protection Impact Assessments (DPIA) designed to help systematically analyse, identify and minimise any risks are an important tool to mitigate risk and determine whether or not the level of risk is acceptable. By building on existing industry practices, relying on existing ethical and legal rules, and avoiding duplication of regulatory requirements, the developments of AI can follow in a responsible and trustworthy manner, while minimising the regulatory compliance to which innovative efforts are subjected.

The use of pre-market access requirements for AI solutions should also be carefully considered. First, as highlighted above, risk could arise after a solution has been put on the market, during the technology’s deployment or usage, rather than at the manufacturing stage. Second, imposing such pre-market assessment on standalone software, which are easy to distribute, should be carefully assessed in terms of proportionality so as to not limit innovation. Software can also be easily updated in order to fix potential concerns, compared to physical products. Thirdly, this kind of regulatory approach requires clear ways of measuring and demonstrating compliance, such as the availability and use of relevant technical standards, the availability of testing protocols to test implementations against those standards, and, even if used on a voluntary basis, notified bodies with the technical experience and bandwidth. Given these limitations, pre-market access requirements should only be considered for very high-risk applications.

Labelling

Overall, EDiMA remains sceptical on the impact of a mandatory labelling scheme on the uptake of trustworthy AI in Europe. An administrative burden to comply with the onerous labelling obligations -- in particular if drafted on the basis of the update Assessment List for Trustworthy AI from the EU High-Level Expert Group on AI -- could significantly outweigh the benefits of such a scheme.

In general, voluntary labelling schemes can play a role in promoting consumer trust, if they are well designed, recognized and sufficiently specific. It remains unclear how this could play a role for AI powered services and products, given the limited amount of information provided by the European Commission at this stage. Indeed, its usefulness would very much depend on the scope of the products considered, their applications, and who the end-user is.

A very general voluntary label may have limited impact, be misleading or could even lead to adverse outcomes if poorly designed. The use of AI may not be immediately obvious or important to the user, but rather works in the background to improve user experience or provide support to the user. Some products may involve very different AI-powered features, which may be hard to reflect on a single label, while multiple labels will impact the user experience and lead to label fatigue, desensitizing the user to the intended message. It is also unclear whether the label will be product specific or could apply to an entire entity.

Voluntary labelling would also require extensive both pre-labelling assessment and post market surveillance schemes, which raises questions linked to proportionality. Voluntary labeling should therefore be considered in specific contexts, focusing on areas where its user trust has been a proven deterrent to adoption.

Human oversight

EDiMA believes that the way in which a particular system is designed to function is an essential element when determining what type of human oversight is needed and appropriate for a certain AI application - or if it is needed at all. This is in line with the Commission's White Paper on AI, which puts forward different approaches to human oversight depending on the "intended use of the system and the effects that the use could have".

It is important to fully understand the direct and indirect impact that human oversight can have in a certain context, while keeping in mind the operational expectation to put in place AI systems. While certain forms of oversight may be critical in one AI application because of its consequential impact, it can be detrimental in others as it could lead to a reduction of quality and accuracy in its output, or even privacy-unfriendly results that fail to take into account GDPR principles such as data minimisation. In this context, when considering human oversight and/or sharing of data sets more broadly, it will be important to ensure that innovative techniques that have significant privacy benefits such as the use on-device processing aimed at avoiding central logging of data (so-called federated learning) are not undermined, but rather encouraged.

Safety and liability implications of AI, IoT and robotics

The safety of users is a top priority for EDiMA members. It is therefore essential that consumers who make use of products and services that rely on AI feel confident that they will find a comprehensive offer of safe and reliable digital services in the online environment.

In this context, EDiMA believes that current product safety legislation at EU level continues to be adequate in light of new technological developments, including AI and other emerging digital advancements such as IoT. As the EU product safety legislation is being separately reviewed at the moment[1] it is essential to ensure coherence so that any new AI regulatory initiative does not overlap with these reviews and ultimately lead to double or contradictory requirements. This is also true of current national liability rules which - through existing laws on damages in contractual and extra-contractual liability rules - address any potential harmful effects of the operation of emerging digital technologies including AI.

As with many other products and services offered in the market that have complex supply chains, in most cases AI products and services can be associated with a person or business entity that can be held liable. There is nothing inherently different about AI that would make it more difficult for a person to find a party liable and seek compensation accordingly.
It is essential that any change to the EU's liability rules maintain the values of the existing framework, which have been provided a solid and stable foundation that incentivised investment, innovation and smart risk-taking for years.

[1] I.e. the General Product Safety Directive (GPSD) and Product Liability Directive (PLD) in particular.