



COMMERZBANK

A risk-based approach to Artificial Intelligence

15 February 2023

Commerzbank's Position on the European Artificial Intelligence Act



COMMERZBANK

Key Messages

- Ad-hoc analysis using ML techniques is out-of-scope of the AI Act.
- A symbiosis between human and machine for labelling tasks, feedback loops and decision making increases the efficiency and effectiveness of a process while at the same time risks can be mitigated.
- Most of the risks related to the use of AI models and systems in financial services are not new and are well-known from successfully handling traditional models in the past. Banks can leverage on existing structures.
- Only where a "yes or no"-loan-granting-decision is involved, the system should be regarded as high-risk as per Annex III, 5(b).
- Certificates might not yield the intended benefits but can quickly become costly and create redundant burdens for banks. It is important to consider the already existing and on-going credit scoring model supervision by competent authorities in the banking sector as equivalent to certifications.
- It must be determined how banks can use certificates. Being able to rely on these is especially important for efficiently implementing state-of-the-art general purpose AI.
- The right amount and form of explanation cannot be determined without specifying the addressee and context. Consumers need to receive the appropriate amount of information in a comprehensible manner allowing them to scrutinize decisions that have been made.
- However, valid concerns speak against disclosing detailed information on AI systems. Especially in the context of fighting fraud disclosing details about the tools used to find fraudulent activities might help circumventing them.
- Due to the entanglements of Artificial Intelligence with various other regulations like GDPR we call for a coherent harmonization of rules.

A. Abstract

Artificial Intelligence (AI) and Machine Learning (ML) are vital for the future of the banking industry to overcome the various challenges of the digital age and the additional (cyber-) risks arising from it. Implementing AI and ML in the highly regulated banking industry is complex. Here, a risk-based approach helps balancing the benefits of this technology against the regulatory hurdles surrounding it.

This position paper shall summarize the most important aspects of the AI Act from the banking perspective. For further details please refer to Commerzbank's white paper which outlines Commerzbank's risk-based approach to ML governance and aims to demystify AI and ML for what it really is: controllable and non-magical!

B. Current Regulatory Environment

At the moment, the European Union (EU) is in the middle of the legislative process to establish harmonized rules on AI, hereinafter: "AI Act"¹. The AI Act aims to regulate the AI components of IT systems regardless of the economic sector they are used in. The regulation is pioneering worldwide and will provide a quality seal for trustworthy AI made and used in Europe.

C. Definitions

Our definition of AI is in alignment with the definition of an AI system according to the EU AI Act (cf. Article 3 and p. 6 (6, 6a-b)). Based on this definition we derive that ad-hoc analysis using ML techniques is out-of-scope of the AI Act.

¹ All further references are based on the General approach of the Council of the European Union as of 6th December 2022 "[Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence \(Artificial Intelligence Act\) and amending certain Union legislative acts](#)".

In early December 2022 the Council of the European Union adopted its common position. Once the European Parliament adopts its own position the 'trilogues' between European Council, Parliament and Commission can be entered. As of today, this is expected for the second quarter of 2023.



COMMERZBANK

D. Risk-Based Approach

AI innovation leads to shifts in risk and return trade-offs, and successful AI implementations must always take the costs of risk mitigation into account. The choice of model depends on the specifics of the situation:

- A high level of quality and a large spectrum of data is a prerequisite and improves traditional and ML models alike.
- In some cases, simpler data driven approaches or “classical” statistical methods work similarly well as more advanced AI models, but without some of the associated risks. Nevertheless, describing complex, non-linear relationships usually requires ML methods to be used.
- By entering a symbiosis between human and machine for labelling tasks, feedback loops and decision making, the efficiency and effectiveness of a process can be increased while at the same time risks can be mitigated.

Hence, Commerzbank pursues a risk-based approach. Systems containing AI are categorized into AI Risk Classes depending on a variety of factors such as model complexity, impact range and business criticality.

Creditworthiness

The following banking function is labelled as a high-risk AI System according to subparagraph 5(b) of Annex III of the European Council’s proposal:

“AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score (...)”

While the intention is clear, i.e., preserving citizens ability to have access to loans, the expression “evaluate the creditworthiness” leaves room for interpretation.

It is not this paper’s intention to lower the standards, but to argue for a realistic, effective, and adequate implementation. We take it for granted that this is about situations where there is specifically a “yes or no”-loan-granting-decision involved as one of the aims of the AI Act is to protect customers’ ability to get a loan. Hence, we conclude that lifecycle applications² are exactly what is remedied for with the exemption of AI components that have a “purely accessory” character (Article 6).

Machine Learning Governance

Risk management at Commerzbank follows the principle of “three lines of defence”. Most of the risks related to the use of AI models and systems in financial services are not new but already well-known from successfully handling traditional models in the past. Hence, banks can lever on existing internal knowledge and structures.

E. Certificates

Certificates and CE markings should increase trust in the solutions offered and consequently support and foster innovation and AI investments especially for small and medium-sized companies. While the intention is clearly in support of innovation, the concept has limitations:

- What *exactly* is certified? It can only ever be a point-in-time snapshot or backwards oriented. It may not hold true beyond a certain time horizon.
- Circumstances like changes in the data could fundamentally change the explanatory power and up-to-dateness of a certification. Hence, it would be reasonable to certify the whole model lifecycle process. Yet a complex certificate cannot easily be understood by humans and only a

² During a customer’s lifecycle we might process information related to a person’s creditworthiness (such as lifecycle applications like information gathering or classification tools).



COMMERZBANK

human understandable certificate is effective and can establish trust as intended.

- Third-party certification only testifies on how well the model works in the designed environment while it might work less efficient or unintended in other environments. It could be mistaken as a “carte blanche” to use the model anywhere.
- It is questionable to what extent banks can rely on this testimony or will need to perform these kinds of evaluations themselves since responsibility can never be outsourced. Hence, it must be determined how banks can use certificates. Being able to rely on these is especially important for efficiently implementing state-of-the-art general purpose AI.

Certificates might not yield the intended benefits but can quickly become costly and create redundant burdens for banks. Credit scoring models are already audited by competent authorities. This should be regarded as a certificate of high quality itself. Consequently, a harmonized standardization should highlight that this ongoing supervision fully meets certification requirements.

F. Trustworthy and Responsible AI

Here, we want to focus on those three aspects with the most vivid debate in public with most room for interpretation: Transparency, explainability and fairness.

Transparency

Transparency is about being clear, open, and honest about how and why a person’s data is being used. (cf. Article 13). Information requirements on data gathering and usage as well as on automated processing are already laid out in Articles 13, 14 and 22 of the GDPR³.

³ The data subject shall “have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her

Explainability / Interpretability

The right amount and form of explanation cannot be determined without specifying the addressee and context. Consumers need to receive the appropriate amount of information in a comprehensible manner allowing them to scrutinize decisions that have been made. On the other hand, for example for the purpose of audits, material needs to be complete and kept for certain time periods.

Valid concerns speak against disclosing detailed information on AI systems. Especially in the context of fighting fraud disclosing details about the tools used to find fraudulent activities might help circumventing them. This is obviously not a desirable result. Furthermore, business know-how has often been developed in-house with large investments in both time and money. Making this public can lead to a significant loss of intellectual property and confidential business logics and result into disadvantages over competitors.

Fairness

Discrimination can be manifold and regardless of the designers’ intentions it can enter the algorithm at various points in the modelling phase. It needs to be carefully monitored in all stages of the model lifecycle with appropriate internal governance in place.

When a potentially discriminating feature is not recorded or later deleted from the dataset, it is still mathematically possible that the model might discriminate around this feature. This is the case when the discriminating feature is correlated to the output⁴.

AI and ML model creation is an iterative process. Of course, transparency and fairness should be a pre-requisite. Nevertheless, lots of corresponding validation steps can only be completed during or after the modelling

or similarly significantly affects him or her.” (Article 22 para 1 GDPR).

⁴ E.g. Amazon Hiring Tool: IIF [Bias and ethical implications in machine learning](#) (p. 10)



COMMERZBANK

process. To ask for final fairness validations from the start without blocking innovation from the beginning is thus neither feasible nor meaningful. Robust and trustworthy AI systems can also help to overcome subconscious biases humans might have as AI yields repeatable and traceable results.

G. Relationship with other laws and jurisdictions

From a regulatory point of view the development, training, evaluation, and deployment of AI systems needs to adhere to various requirements⁵ in addition to the upcoming AI Act.

Other jurisdictions in the world are about to adopt AI-related regulation with different angles and priorities while some argue that most AI topics are already covered sufficiently by existing regulation and might only require to be detailed accordingly. These developments pose the risk of market fragmentation and put challenges to multinational organizations in general.

H. Conclusion

We appreciate the general orientation the AI Act provides – an approach to a quality seal of AI made or used in Europe. Definitions are state-of-the-art and a risk-based approach is suitable. Since the Act shall apply for all various industries some banking specifics are omitted. This is especially important when considering the fine line between lifecycle customer management and credit scoring as well as the drawbacks additional certifications have in an already externally audited field of expertise like credit scoring. Here it is important to consider the on-going model supervision by the competent authorities as equivalent to certifications.

It is not this paper's intention to lower the standards, but to point to the already existing and effective regulation in the banking industry and to argue for a realistic, effective, and adequate implementation of the new requirements. Due to the entanglements of Artificial Intelligence with various other regulations as described above we call for a coherent harmonization of rules.

⁵ Among these are: Directive 2013/36/EU, General Data Protection Regulation, European Data Act, European Data Governance Act, BAIT; MaRisk, EBA Guidelines on ICT, DORA,

MiFiD and at least indirectly through many regulations like laws on consumer protection, equality, antidiscrimination etc.



COMMERZBANK

Commerzbank AG

Zentrale
Kaiserplatz
Frankfurt am Main
www.commerzbank.de

Postanschrift
60261 Frankfurt am Main
Tel. + 49 69 136-20
info@commerzbank.com

Big Data & Advanced Analytics

Julia Sterling
Julia.sterling@commerzbank.com

Thomas Stadje
Thomas.stadje@commerzbank.com

Commerzbank's division "Big Data & Advanced Analytics" is the Center of Competence for all AI and ML related matters which range from in-house AI modelling and implementation to advisory tasks.

