Principles on the use of Artificial Intelligence (AI) and Machine Learning (ML)

Principles of Commerzbank AG[1] and domestic branches regarding their employees

Preamble

Artificial Intelligence (AI) and Machine Learning (ML) are increasingly finding their way into our daily life and into Commerzbank's working environment. AI and ML open up a wide range of opportunities to serve our clients in a better way, to make banking processes faster, better and more efficient as well as to reduce losses due to risks, fraud and financial crime. At the same time, the implementation of AI and ML models changes the working processes of our employees. The use of AI is intended to support employees in their work and to make processes more efficiently. The goal is also to ensure that employees are empowered to deal with the changes related to AI.

For these reasons it is important that the employer and the "Konzernbetriebsrat" (i.e. Group Works Council) establish a common view regarding the use of Artificial Intelligence and Machine Learning in Commerzbank. The common goal is to take worries and fears regarding this sensitive topic seriously, to integrate the employees and their know-how in new processes and changed working conditions, to make the use of AI and ML transparent and to explain it in an appropriate manner. On that basis, the advantages of these technologies can be levered and at the same time errors and misuse can be avoided.

The terms „Artificial Intelligence" and „Machine Learning" are not being used consistently in general. Traditional programming is based on rules set by the programmer. In contrast, the principles agreed here use the term AI system for the implementation of models that „learn" such rules from the data available to them with a significant degree of autonomy and flexibility.[2]

Principles

The personal rights of our employees are respected and protected in the context of the use of Artificial Intelligence and Machine Learning in our group, in accordance with existing laws. By this we mean:

- As a group, we are aware that AI systems can affect the rights and opportunities of others.
- AI systems must take into account the fundamental rights of those affected, as well as applicable laws and democratic values.
- AI systems must process employee data in a comprehensible way.
- The strategic orientation of the group in the field of AI and ML as well as the relevant concepts and terms are presented and explained to the employees in a transparent manner.
- Opportunities and risks are addressed and weighed when selecting models and implementing AI systems.

[1] This applies to Commerzbank AG in Germany as well as to foreign branches and DTCs.

[2] For details pls. refer to glossary

- The use of AI systems as well as their quality standards are aligned in such a way that discrimination is avoided.
    - We ensure that AI systems are used in a trustworthy and responsible manner.
    - In particular, this includes fairness, transparency, robustness, security, data protection compliance and human control.
- When people interact with chatbots, talkbots and deep fakes, transparency about the interaction with an AI system is created on the front end in a timely, clear and comprehensible manner.
- In processes with the support of Artificial Intelligence, it must always be clear and comprehensible to employees which rights and obligations exist (e.g. intervention options and mandatory testing).
- "Mitbestimmungsrechte des Betriebsrats" (works council's rights of co-determination) and existing "Betriebsvereinbarungen" (company agreements) also apply to the use of Artificial Intelligence and are not restricted by these principles[3]. In particular, the provisions of the IT-KBV framework apply.
- The use of AI systems in personnel-relevant processes (HR processes) is considered separately in the context of the existing co-determination rights of Works Council outside the present principles.

Our employees must not become glassy through the use of AI systems. This means:

- Any manipulative, exploitative and social control practices[4] are not used in AI systems.
- AI systems are not used to analyze or manipulate emotions, personality traits or mental health of our employees. AI systems designed to improve the well-being of employees are only used if permitted by company agreement. No AI systems are used for the purpose of ascribing character traits.
- Article 9 (1) of GDPR also applies to AI systems*: „Processing of personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation shall be prohibited".* [5] Exceptions to this rule are provided for in GDPR Article 9 (2).

Commerzbank and the employee representatives will continue to exchange views on the use of Artificial Intelligence and Machine Learning on a regular basis. The focus of this joint dialogue is in particular on the continuously developing use of AI and ML in the bank, the review and further development of the existing principles as well as their operationalization.

Glossary

term	explanation
Artificial Intelligence (AI)	Artificial Intelligence refers to the creation of an algorithm that uses data to model certain aspects of the world. The model is then applied to new data in order to generate results in the form of e.g. content, predictions, classifications or recommendations. The aim is to support human activities or ultimately to make certain decisions.
AI system	"AI system" means a machine-based system that can generate predictions, recommendations or decisions that affect the physical or virtual environment

[3] According to Data Protection Manual (SEC-32002-EN – Valid from 01. June 2023 | Version 12.0 Chapter 21), the following applies: „If personal data is processed by employees, the responsible workers council must be consulted." In addition, all agreements of the written framework continue to apply to the use of Artificial Intelligence, in particular the principles of data protection and IT security. .

[4] For example, „Social Scoring" falls under this category. „Social Scoring" means the assessment or classification of individuals based on their social behavior, socio-economic status, acquaintances or predicted personality traits.

[5] This does not affect verification systems that only compare a person's biometric data with their previously provided biometric data. It also excludes systems that are used to identify natural persons remotely with the sole purpose of confirming whether a particular natural person is allowed to gain access to a service, device or location.

	with varying degrees of autonomy. Within the framework of these principles we use the term AI system for software that uses Machine Learning, more precisely „non-parametric“ algorithms with an existing degree of autonomy.
Machine Learning (ML)	Traditional programming codifies rules (e.g. if/then conditions). It freezes the modeled reality. By contrast, ML aims to identify („learn“) these rules, which can be too complex for traditional programming. ML models represent statistic input-output relationships and are very well suited for describing complex and non-linear relationships, allowing them to respond to new data or changes in reality. Often the result provides probabilities for various possible outputs that require acceptance thresholds to use the respective output (e.g. for a „yes or no“ decision). If the mathematical function that is to be learnt is known or defined in advance, this is referred to as „parametric algorithms". In this case, „only“ the parameters of the specified function are learnt from the data. Examples of „parametric algorithms“ are linear regressions or logistic regressions. In contrast, there are „non-parametric algorithms“. These do not make strong assumptions about the form of the function and can thus learn any functional form and distribution from the training data. Examples of „non-parametric algorithms“ are neural networks or ensemble methods (e. g. Gradient Boosting, Random Forest).
Fundamental Rights	The Universal Declaration of Human Rights of the United Nations applies internationally. Furthermore, the concept of the Fundamental Rights covers the values set out in Article 2 of the Treaty on European Union (TEU) on which the European Union is founded: Respect for human dignity, freedom, democracy, equality, the rule of law and respect for human rights, including the rights of persons belonging to minorities. The Charter of Fundamental Rights of the European Union also applies. These Fundamental Rights are supplemented by national regulations such as the “Grundgesetz” (i.e. Basic Law in Germany). Since these are fundamental rights of individuals, they must be protected to a particular extent.
Fairness	AI systems should not lead to unjustified, adverse treatment of individuals, a so-called discrimination. Discrimination means that someone is degraded or treated as morally inferior because of a socially prominent trait. Fairness prevails where there is a kind of equality in which discrimination is largely prevented, monitored and mitigated. Discrimination can arise at different points both in the modeling process itself, during the model life cycle or through the environment used.
Transparency	Transparency primarily asks whether the basic functionality of an AI system is sufficiently comprehensible for users, those affected, experts, management and all external stakeholders, and whether the results of the system can be reproduced and justified if necessary. It covers aspects of traceability, reproducibility and explainability.
Robustness	Robustness addresses the risks that arise when input data are corrupted or manipulated, but for which data the ML component is designed to accurately process it (for example, wrong characters in a word).

Security	Security includes both functional security features and protection against attacks and manipulations vs. the AI system.
Data protection conformity	This dimension refers to the protection of sensitive data in the context of the development and operation of an AI system. This concerns both the protection of personal data and business secrets. Here we refer to the existing data protection policy and the data protection manual as well as the SECAM classification of an IT system, which also includes data protection-relevant aspects.
Human control	AI systems are usually used in situations where there is a form of interaction with the outside world and typically also an interaction with humans. The possible spectrum ranges from full human control („Human Control“) to automated decision-making („Human-out-of-the-loop“). In between, a distinction is made between „Human-in-the-loop“ and „human-on-the-loop“. In the former, the person is heavily involved in most decisions and can take over the process at any time, while in the latter, the person has a monitoring function and the system takes most decisions.
Deep Fake	Deep Fake generally means manipulated or synthetic audio, image, or video content that falsely appear to be authentic or truthful. They contain images of people who seem to say or do things that they did not say or do, but were the product of AI.