FINMA Guidance 08/2024 (link)

Governance and risk management when using artificial intelligence (1/3)

Topics	Pillar 1: Governance	Pillar 2: Inventory & risk classification
FINMA's findings	- Supervised institutions primarily focus on data protection risks and not on model risks associated with AI. - The development of AI applications is often decentralised. - In the case of externally purchased applications, it is sometimes difficult to determine whether AI is included.	- Some supervised institutions defined AI narrowly - Difficulty for some supervised institutions to ensure the completeness of the inventory. - Lack of criteria to identify AI applications that present a risk
FINMA's expectations	- Centrally managed inventory for AI applications - Responsibilities and accountabilities (for the development and use phases) must be clearly defined. - Setting up rules for model testing, documentation standards, and broad training measures - In the case of outsourcing, contractual clauses governing responsibilities and liability issues of the provider	- A sufficiently broad and uniform definition of AI - Establishment of criteria to identify significant AI applications and specific risks requiring special attention
Concrete implementation (examples from OA practice)	Implement an internal directive that allocates responsibilities (to avoid: projects managed solely by the first line or that emerge in a decentralised manner), *but important to note* what FINMA does not say: the authority does not prohibit decision-making through AI (however a "human" must assume ultimate responsibility) → acceptability of automated individual decisions (Art. 21 of the Swiss Data Protection Act) Understanding the services provided by third parties + contractual commitments to be obtained from third parties (even if the market is concentrated with a small number of providers) → similar issues as those arising in the context of outsourcing projects	Checklist and heat map to document, for each use case, (i) model risks (robustness, correctness, bias, stability, and explainability) of AI and (ii), where applicable, the contractual commitments made by third-party providers. Inventory with a risk-based classification

FINMA Guidance 08/2024 (link)

Governance and risk management when using artificial intelligence (2/3)

Topics	Pillar 3: Data quality	Pillar 4: Tests & ongoing monitoring	Pillar 5: Documentation
FINMA's findings	• Not all supervised institutions have defined rules and processes to ensure data quality in AI applications.	• Weaknesses identified in the planning and implementation of tests and controls • Few specific performance indicators are defined in advance.	• Some supervised institutions do not have directives to document the use of AI. • Documentation which is incomplete, insufficiently detailed, and not tailored for the users of the application
FINMA's expectations	• Establish internal rules/directives to ensure the completeness, correctness, integrity, and accessibility of the data used	• Implementation of testing processes to verify AI models, and to ensure that the applications achieve the intended objectives. • Conducting regular checks of AI outputs	• Provide detailed documentation for important applications covering: the objectives of the application, its reliability, risks, data selection, and data quality.
Concrete implementation (examples from OA practice)	Internal process for the quality control of the input data Implicit scepticism of the regulator towards the use of LLMs (due to the very practical difficulty of data quality control) → risk that the deployment of LLMs may be subject to regulatory limits in the future?	Definition of KPIs. *Ex post* controls to address the phenomenon of model/data drift Audit process (if necessary, by a third-party expert)	Documentation of the applications used: (i) purpose of the applications, (ii) selection and preparation of data, (iii) selection of models, (iv) KPIs, (v) tests and controls, and (vi) fallback solutions

FINMA Guidance 08/2024 (link)

Governance and risk management when using artificial intelligence (3/3)

Topics	Pillar 6: Explainability	Pillar 7: Independent review
FINMA's findings	• The results of AI models are often not explainable and not reproducible, which limits the ability to conduct critical assessment.	• Independent verification processes for the development of AI models are rarely implemented.
FINMA's expectations	• Ensure that the results of the models are understandable to stakeholders, whether they are for example investors, clients, or employees. • Understand the mechanisms behind the models to ensure their plausibility and robustness.	• For material applications, implement an independent review covering the entire development cycle to obtain objective opinions and identify risks.
Concrete implementation (examples from OA practice)	Due diligence of AI applications → the (challenging!) goal should be the reproducibility of the output to be able to understand its origin / sensitivity analysis / indication of sources in *Retrieval Augmented Generation (RAG)* projects Also important for defending against third-party liability claims → "premonitory" case law: ATF 4A_301/2023 (in the case of liquidation of a client's positions with a negative balance, the bank must prove these losses)	Functional separation (Chinese walls?) between developers and those responsible for the review Involvement of third-party experts at both technical and legal levels, particularly for defining KPIs, output control, and audits