

Just in Time

Big Data and Artificial Intelligence: *Principles for the Use of Algorithms in Decision-making Processes*

Sep 2021

Executive Summary

In the financial market, there is great interest in the use of big data and artificial intelligence (BDAI). However, the use of BDAI is also associated with risks that must be monitored:

- **Artificial intelligence as a combination of machine learning and big data**
The technical definition of the term “artificial intelligence” (AI) is the combination of big data, computing resources and machine learning (ML).
- **Currently it is impossible to make a clear distinction between AI and traditional processes**
It is impossible to make a clear distinction between AI on the one hand and traditional statistical processes and the algorithms that are used within this context on the other. Further developing the existing definition of AI is one of the challenges faced by supervisors, regulators and standard-setters.
- **Considerable complexity, short recalibration cycles and a high level of automation**
Three key features which characterize modern artificial intelligence methods and applications, and which play a role when observing risks are:
 - I. Machine learning algorithms are often more complex than those used in traditional statistical processes
 - II. The combination of self-learning algorithms, and mass data leads to ever-shorter recalibration cycles for models and algorithms
 - III. Algorithms are also increasingly being used in some cases to automate non standardized processes and decisions that are made quickly and in large numbers
- **Minimum requirements for the use of AI as a basis for discussion**
The general principles for the use of algorithms in decision-making processes at financial entities:
 - I. constitute preliminary ideas for minimum supervisory requirements relating to the use of artificial intelligence and form the basis for discussions with various stakeholders
 - II. can already serve as guidance for the entities under BaFin's supervision.



At a Glance

01	<u>Conceptual Framework</u>	4
02	<u>Key Principles</u>	7
03	<u>Specific Principles for the Development Phase</u>	10
04	<u>Specific Principles for the Application Phase</u>	15
05	<u>Inclusion of Principles in International Regulatory Projects</u>	20
06	<u>Final Remarks</u>	22



01

Conceptual Framework



Conceptual Framework 1/2

1

Considering the Process as a Whole

- **Whether algorithms yield usable results depend on how supervised entities incorporate these algorithms into decision-making processes.** An algorithm that is suitable for a particular context may yield unusable results in another situation.
- Furthermore, the results of an algorithm depend on data availability and data quality.

2

No General Approval Process for Algorithms

- **Supervisor examines and raises objections to such processes in a risk-oriented manner** as and where needed.
- In justified exceptional cases the supervisor examines whether the algorithmic processes that are used are appropriate in terms of methodology, calibration and validation, among other things.

Conceptual Framework 2/2

3

Risk-oriented, Proportional and Technology-neutral

The fundamental supervisory and regulatory principle of “same business, same risk, same rules” means that a **risk-oriented, proportional and technology-neutral approach must be taken for the supervision of algorithmic decision-making processes.**

- Accordingly, **more intensive supervision is in order if (additional) risks are associated with the use of an algorithm in decision-making processes.** One typical risk is the significant scalability of processes and thus potential errors.

4

Existing Rules are Supplemented, Specified and Further Developed

These principles are intended to integrate and specify the existing regulations and administrative practices in place and **constituted preliminary ideas for minimum supervisory requirements relating to the use of artificial intelligence.**

- Many of these principles **bring about changes in certain areas of principles-based and technology-neutral regulation.**
- The principles in this publication do not rule out the fact that certain regulated activities may already be subject to stricter regulations or administrative practices. In such cases, compliance with these rules takes precedence.

02

Key Principles



Key Principles 1/2

1

Clear Management Responsibility

Senior management is responsible for all significant business decisions, even if they are based on algorithms.

- Senior management must have sufficient technical expertise.
- The business-wide strategy for using algorithm-based decision-making processes should be reflected in the IT strategy.
- There must also be staff with the necessary technical knowledge in the independent control functions

2

Appropriate Risk and Outsourcing Management

Senior management is also responsible for establishing a risk management system and if applications are used by a service provider, senior management must set up an effective outsourcing management system.

- A “model risk management framework” should be established at companies to consider all algorithm-based decision-making processes and interdependencies.
- Risk mitigation measures processes should be targeted and applied precisely where risks originate.
- Measures to minimize cyber security risks should also be adapted if required.

Key Principles 2/2

3

Preventing Bias

- Bias must be prevented in order to be able to **reach business decisions that are not based on systematically distorted results**;
- Bias must be prevented in order to rule out bias-based systematic discrimination of certain groups of customers and thus **rule out any resulting reputational risks**.
- Identify the risk of bias where it may occur, considering the root cause, and to analyze this risk and either eliminate or at least mitigate this risk.

4

Ruling out Types of Differentiation that are Prohibited by Law

In the case of certain financial services, the law stipulates that certain characteristics may not be considered for differentiation purposes. If conditions are systematically set out based on such characteristics, there is a risk of discrimination. This would be associated with increased reputational risks and, in some cases, legal risks.

- Companies should therefore establish **(statistical) verification processes to rule out discrimination**.

03

Specific Principles for the Development Phase

Use Case: Additional Information Analysis for Credit Ratings



Specific Principles for the Development Phase 1/4

1

Data Strategy and Data Governance

- Depending on the application and features of the algorithm, data must be used in **sufficient quality** and **quantity**.
- Companies must have a **verifiable data strategy** which guarantees the continuous provision of data and must be implemented in a data governance system defining **responsibilities**.

2

Compliance with Data Protection Requirements

- **Compliance** with the applicable **data protection requirements** must always be ensured.
- Data protection requirements for the use of data should already be considered when planning algorithmic decision-making processes.
- Disclosure requirements vis-à-vis data subjects must also be observed.

Specific Principles for the Development Phase 2/4

3

Ensuring Accurate, Robust and Reproducible Results

- The ultimate objective is to ensure **accurate** and **robust results**.
- It should also be possible to reproduce the results of an algorithm. By ensuring that, the results can be understood and verified at least to a certain degree **by individuals within the company** and by **external unit**.
- In addition, caution and precision are ensured for the selection of the algorithm as well as the calibration and documentation process.

4

Documentation to Ensure Clarity for both Internal And External Parties

Sufficient documentation is required in order to ensure that algorithms and the underlying models can be verified:

- **The selection of the model must be documented.** It is essential to determine whether the model is suitable for specific application using statistical considerations and taking into account the complexity of the model and its interpretability and verifiability.
- **The model calibration and training must be documented.**
- **The model validation must be described.** In certain circumstances, a clear distinction cannot be made for the documentation of the selection, calibration and validation of the model.
 - In some instances, the quality of a model can only be determined after the initial calibration and validation of the model
 - Furthermore, it is only possible to select a model by comparing different models.

Specific Principles for the Development Phase 3/4

5

Appropriate Validation Processes

- Every algorithm should go through an appropriate **validation process** by an **independent function** prior to being included in operations.
- It is also necessary to determine and document the intervals at which an algorithm must undergo another validation and the **factors** that will lead to the **ad hoc validation**. Such factors include:
 - A systematic **change in input data**;
 - **External** (macroeconomic) **shocks**;
 - Change to the **legal** requirements;
 - Feedback from the output phase such as a **threshold being crossed**.

6

Using Relevant Data for Calibration and Validation Purposes

- A calibration or validation with unsuitable or erroneous data can systematically **affect the performance** of the algorithm in live operations.
- The data must be relevant and representative for the application in question.
- Depending on the scope and riskiness of the decision for which an algorithm is used, various measures should be taken to ensure that the calibration and validation can be subsequently understood and verified.
- In **highly risk-sensitive** decision-making processes, the data that is used for this purpose should be saved and stored whereas in **less risk-sensitive** decision-making processes, the selection process and the structure of the data should be documented at a minimum.

Specific Principles for the Development Phase 4/4

Use Case: Additional Information Analysis for Credit Ratings



→ A credit institution has established a process to determine **credit ratings** of companies, **using natural language processing (NLP)**.

This new process involves **random forest approaches of machine learning**, which are randomly generated decision trees with optimized prediction quality.

- The insights that are gained in this way supplement the established credit rating based on quantitative corporate information. If there is a **significant discrepancy** between the established approach and the credit rating based on the ML approach, a credit analyst must reach the final decision.
- The analyst compares the ML result with an **independent expert assessment** (“**putting the human in the loop**“ principle).
- The process must document the results from the training process in a suitable manner for the expert concerned, in line with the principle of “**documentation to ensure clarity for both internal and external parties**”.

1 Putting the human in the loop*

2 Documentation to ensure clarity for both internal and external parties

* For further details see the next slides.

04

Specific Principles for the Application Phase

Use Case: Sanction Screening in the Context of Money Laundering Detection

Use Case: Fund Management



Specific Principles for the Application Phase 1/4

1

Putting the Human in the Loop

Employee should be involved in the interpretation and the use of algorithm results for making decisions. This involvement should depend on how mission-critical the decision-making process is, and the risks entailed. Moreover, their involvement should bring real benefits and should not be limited to the mere approval of every algorithm decision.

2

In-depth Approval and Feedback Process

All the situations involving a more in depth-approval process should be **defined in advance in a risk-oriented manner**.

This form could be, for example, a **threshold-based process**:

- if the first threshold is crossed, the approval should be granted by an individual.
- If the second threshold is crossed, the approval should be granted only after an additional review, to determine if there are some peculiarities in the input data.
- If the final threshold is crossed, there should be a signal for examining the model and the interruption of the process.

Such way reduce the risk of erroneous decision in algorithm decision-making process **and improve the results in the long term**.

Specific Principles for the Application Phase 2/4

3

Establishing Contingency Measures

There should be implemented measures that allow business operations to **continue to run if problems arise** in algorithm-based decision-making processes.

4

Ongoing Validation, Overall Evaluation and Appropriate Adjustment

Algorithms must be ongoing validated in order to assess functionality, check for any discrepancies and make adjustment if necessary.

- In order to ensure an independent evaluation of functionality and risks of the algorithms, an internal or external audit **should be performed to examine the regular evaluation and adjustment process.**

Specific Principles for the Application Phase 3/4

Use Case: Sanction Screening in the Context of Money Laundering Detection

1 Putting the human in the loop

2 In-depth approval and feedback process



→ In the money laundering detection context, it is necessary to check and to identify if transactions violate any embargoes. These tasks are carried out by an algorithm and a staff member. If there are critical points (such as discrepancies between the algorithm and employee results) a second individual is involved.

- In such way, effective and risk-sensitive human involvement is still in place and the **putting the human in the loop** principle is ensured.
- the third checks that take place when the algorithm and the employee do not reach same results can be seen in the **in-depth approval and feedback process** principle.

Specific Principles for the Application Phase 3/4

Use Case: Fund Management

1

Ongoing validation,
overall evaluation and
appropriate adjustments

2

Interpretation and use of
algorithmic results for
reaching decisions

3

In -depth approval and
feedback process



→ Investment/disinvestment decisions are not reached exclusively by a fund manager. Such decisions are taken using also a quantitative multi-factorial model implemented through an algorithm.

- The quantitative model is revised on a regular basis and new metrics are added or existing ones are modified: this follows the principle of **ongoing validation, overall evaluation and appropriate adjustment**.
- In the context of **interpretation and use of algorithmic results for reaching decisions**, the portfolio manager takes into consideration shares that show the great performance in terms of model rating and review the composition in a quarterly basis.
- Finally, in order to mitigate short-term market effects, the portfolio manager can adjust the factors or the strategy weighting, within the **in-depth approval and feedback processes** principle.

05

Inclusion of Principles in International Regulatory Projects



Inclusion of Principles in International Regulatory Projects

These principles constitute preliminary ideas for minimum supervisory requirements relating to the use of artificial intelligence. As a result, **they form the ground for discussions with various stakeholders**, including market participants, members of academia, national or international supervisory authorities and standard-setters.

1

In its Digital Finance Strategy, **EC announced its intention to clarify, by 2024 at the latest, together with the ESAs, whether and how existing financial market regulation should apply to BDAI applications.**

2

Various working groups dealing with the risks and opportunities associated with BDAI had already been established at the ESAs (such as the EBA Task Force on IT, the EIOPA InsurTech Task Force and the ESMA Financial Innovation Standing Committee).

3

Work is also being carried out in various areas in the field of BDAI at national and international standard-setters such as the DIN Committee or the ISO Committee. These committees are expected to work on a clear **definition for artificial intelligence and machine learning**. **Further technical standards** (beyond financial market regulation) for the use of BDAI can also be anticipated.

06

Final Remarks



Final Remarks 1/2

The principles set out in the previous slides and published by BaFin are intended to **promote the responsible use of big data and artificial intelligence (BDAI)** and **facilitate control of the associated risks**.

- BaFin principles are expected to **stimulate discussion with ECB and the ESAs**; furthermore, EC have already developed regulatory and supervisory guidance of AI applications.

The **choice** by the institutions regarding the **incorporation of algorithms in decision-making processes** is decisive for making BDAI-based algorithms usable.

The supervision of algorithmic decision-making processes must adopt a risk-oriented, proportional and technologically neutral approach.

- More intensive supervision is necessary if further risks are associated with the use of an algorithm in decision-making processes.

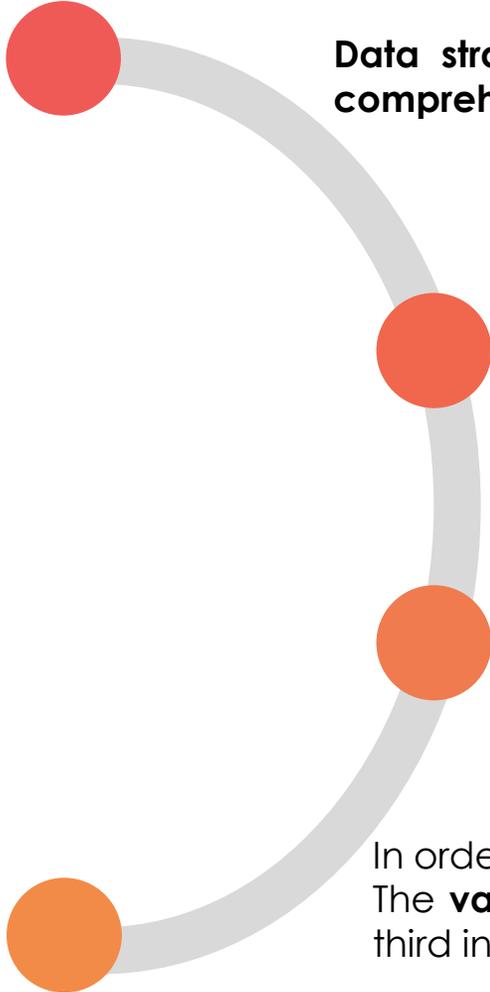
Senior management is **responsible for algorithm-based decisions** and **must develop technical expertise**.

- There must also be staff with the necessary technical knowledge in the independent control functions.

Targeted risk mitigation measures and processes must be applied where risks originates.

- A **model risk management framework** must be implemented in order to consider, holistically, all algorithm-based decision-making processes and interdependencies and to detect possible biases in the processes.

Final Remarks 2/2



Data strategy, data governance and **compliance requirements** must be defined by providing **clear and comprehensive documentation** for both internal and external parties.

The final aim is to obtain **accurate** and **robust results**.
Users should be able to reproduce results in a subsequent test performed by an independent third party.

A process of a more in-depth approval is necessary in order to reduce the risk of erroneous decision reached by an algorithm and individuals could be involved bringing real benefits in algorithm-based decision-making processes

In order to assure functionality and to check for any discrepancies, algorithms must **be ongoing validated**.
The **validation and adjustment process must be examined** involving an internal/external auditor and by a third independent internal/external control function.

Company Profile

Iason is an international firm that consults Financial Institutions on Risk Management. Iason integrates deep industry knowledge with specialised expertise in Market, Liquidity, Funding, Credit and Counterparty Risk, in Organisational Set-Up and in Strategic Planning.

Dario Esposito



Matteo Cecchin



This document was prepared in collaboration with Riccardo Chinello, who at the time was working for Iason Consulting.

This is an Iason creation.

The ideas and the model frameworks described in this presentation are the fruit of the intellectual efforts and of the skills of the people working in Iason. You may not reproduce or transmit any part of this document in any form or by any means, electronic or mechanical, including photocopying and recording, for any purpose without the express written permission of **Iason Consulting Ltd.**

www.iasonltd.com