

NILS GILLES · 29/01/2024

AI Steering & Control Framework

In this article, you can read about the key aspects of such a framework, which focuses on compliance and governance meta-information, technical performance indicators and business metrics.

In today's corporate world, AI and machine learning models are increasingly becoming a critical part of business strategy. To ensure that these models are not only successfully implemented, but also effectively managed and utilized, the introduction of a governance framework is crucial.

In this article, you can read about the key aspects of such a framework, which focuses on compliance and governance meta-information, technical performance indicators and business metrics.

To ensure sustainable success, it is necessary to integrate a continuous feedback loop into the AI lifecycle, which serves as a trigger for re-training and model adjustments based on the metrics mentioned above, but also compares the economic contribution of a model against its costs for operation, infrastructure and maintenance.

In this context, we present our standardized process model for implementation, starting with strategic alignment and definition of metrics through to technical integration, e.g. via a governance layer in your IT landscape.

Meta information

The integration of ML models into business-critical processes requires a high degree of transparency and control. It is therefore essential to capture meta information that provides a clear insight into the use of the models and at the same time ensures that compliance and governance standards are adhered to.

This information can be structured along three dimensions: Utilization, ethics, and reproducibility.

Utilization

Transparency about business processes using ML, their criticality and assessment of the associated risk of ML-supported decisions. Examples of metrics here are the number of affected users, the number of affected systems or the business volume processed with ML support as well as the degree of automation.

Ethics

Monitoring the use and quality of data and ensuring that ML models comply with ethical and compliance guidelines. This includes, for example, the implementation of indicators for the detection of bias in training data or model results as well as warning indicators when using sensitive data during training.

Reproducibility & explainability

Measuring the explainability of models to ensure that decisions are comprehensible. Transparent models not only support compliance requirements, but also promote stakeholder trust.

Metrics such as SHAP on the significance of features for the model statements offer approaches to explainability. Reproducibility, on the other hand, is made possible by the transparency and availability of the model and data versions used during run-time as well as the parameter configuration.

Technical performance indicators

The measurement of ML model performance varies greatly depending on the model class and the specific requirements of the company. The framework should be able to capture technical performance indicators for different model classes, measure response times taking into account use case requirements and keep an eye on technical resource utilization from a cost perspective.

Model accuracy & error rates

Classic measures of model performance. However, the framework should also identify specific error types to better understand weaknesses in the predictions.

Depending on the model class (clustering, language processing, GenAI, prediction/regression, etc.), there are a number of different indicators, some of which are listed in Table 1 as illustrative but not exhaustive examples.

Response times & speed

The time a model needs to make a prediction can be critical. Depending on the area of application, different response time targets must be set. A response time of 60 seconds may be completely sufficient for granting a loan, but this is unacceptable for machine control & autonomous driving, for example.

Use of resources

Monitoring the resource consumption of the models to ensure that they work efficiently and do not consume excessive IT resources.

Increasing costs for individual models over time are valuable information in the context of ML lifecycle management. Very often smaller LLM models are sufficient for the task. gpt3.5-turbo for most tasks provides results that are for business as good as results from OpenAI gpt4, while having 30 times smaller cost.

Business impact

The evaluation of AI & ML use cases is based on their technical and legal feasibility, but their raison d'être in a company is also examined in particular in terms of their economic value contribution.

Return on investment

A clear measurement of how the use of ML models impacts the financial performance of the organization. This includes costs for model development, implementation and maintenance compared to the benefits achieved.

Business process improvement

Analysis of how well the models support and improve existing business processes. This can increase efficiency and lead to cost savings (example KPIs in Table 1).

Customer satisfaction

Monitoring how the models influence customer satisfaction. This can be done by analyzing customer feedback, return rates or other relevant metrics (example KPIs in Table 1).

Ultimately, the economic impact of an AI solution must be continuously reassessed; a one-off calculation of the business case for a model is absolutely necessary in advance, but not sufficient. Fluctuating operating costs, changing business processes or changing model quality have the potential to transform previously unprofitable models into more profitable ones. At the same time, previously well-functioning components could become candidates for re-training or decommissioning.

Table 1 provides an overview of exemplary metrics for the three areas, which must be defined in the respective corporate context and aligned with the AI strategy and thus the overarching business strategy.




 Meta information	 Technical performance indicators	 Business impact
<p>Definition of KPIs to support governance, compliance & operation of the models.</p> <p>Meta information, e.g: Date of last model update Number of users Number of business processes affected Model criticality for business processes Current model version per business process Dependence on other models</p> <p>Indicators for detecting bias</p> <p>Statistics on adjustments to features</p> <p>Data drift indicators</p> <p>Operating costs</p>	<p>Implementation of metrics for monitoring model performance development depending on the model class, e.g.:</p> <p>Classification: Confusion Matrix, Accuracy, Precision, Sensitivity, F1</p> <p>Clustering: Silhouette Score, Davis-Bouldin Index, Calinski-Harabasz Index</p> <p>Forecasting & Prediction: Mean Squared Error, Mean Absolute Error</p> <p>Generative AI: Fréchet Inception Distance (FID), Structural Similarity Index (SSIM), Mean Opinion Score (MOS)</p> <p>Language processing: Perplexity, Word Error Ratio (WER), BLEU (Bilingual Evaluation Understudy) score</p>	<p>The definition for measuring business impact depends on the individual use case.</p> <p>Depending on the use case, typical KPIs:</p> <p>Conversion rates</p> <p>Throughput times of business processes</p> <p>Number of users of a service</p> <p>Customer satisfaction / NPS</p> <p>Migration rates</p> <p>Identified cases of fraud</p> <p>Sales increase for divisions, services & products</p> <p>Cost reduction in the business process</p> <p>Error rates</p> <p>The core challenge at the beginning is to define a baseline for evaluating the impact of a model for its use case.</p> <p>A continuous comparison process then ensures the necessary transparency and defined threshold values for evaluating model performance.</p>

Table 1: Example metrics for performance measurement & governance of ML

Implementing a Five-Step ML Control Framework: Integrating ML into Your IT Ecosystem

As in any KPI-based management framework, it is not the quantity of KPIs measured that is ultimately decisive, but the selection of the right metrics and their balancing.

Based on the AI strategy (and therefore also based on the corporate strategy) as well as on the application context of the use cases, the KPIs that best support the operationalization of the strategic orientation as "leading", or "lagging" indicators must be selected.

As is so often the case, less is often more!

To streamline the processes, an ML governance layer must be integrated into the IT architecture, which centralizes all essential information and makes it available to the cockpit or the ML Ops processes.

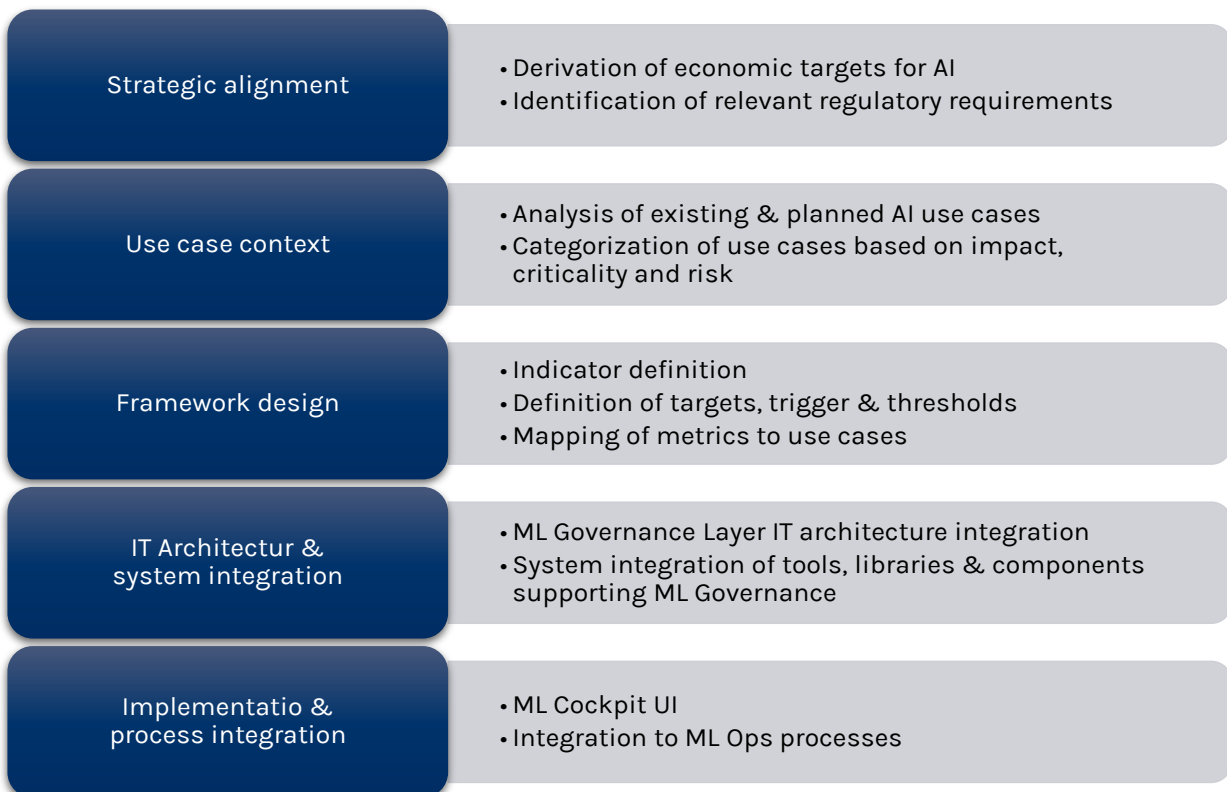


Figure 1: Building blocks for deriving a KPI-based ML control framework

Our additional white papers on "ML Cockpit" and "ML Reproducibility" provide a deep dive into how this framework can be technically implemented in reality and integrated into the technical architecture of your IT.

Feel free to contact us if you would like us to accompany you on your journey to the transparent and efficient use of AI in your company!



Ante Gojsalic
Lead Data Scientist
ante.gojsalic@trusteq.de



Nils Gilles
Head of Data & Analytics
nils.gilles@trusteq.de