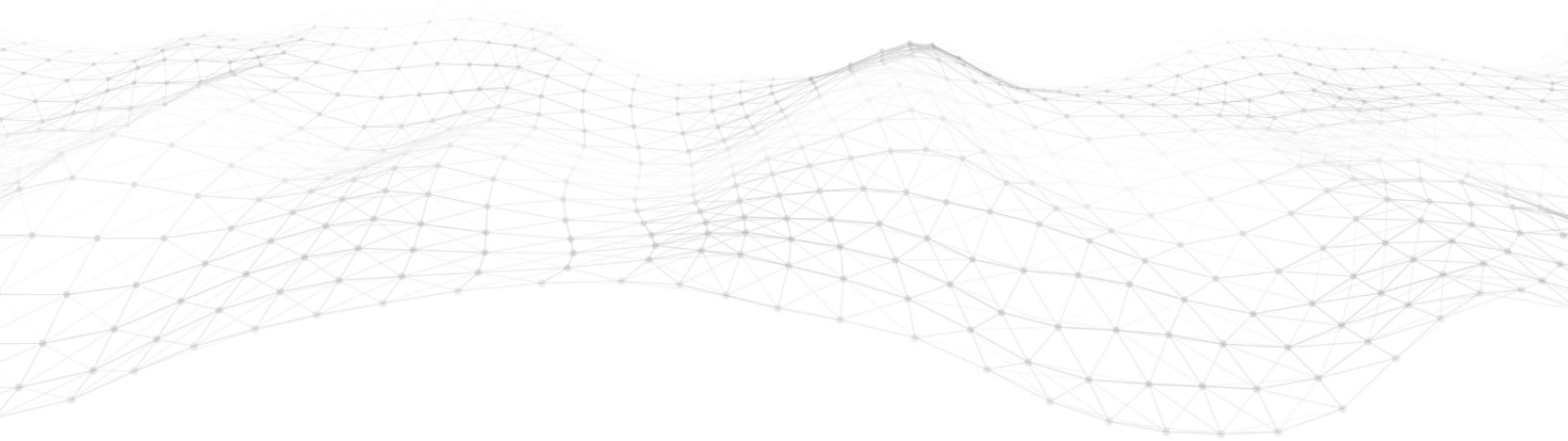
AI Model Governance White Paper

6/27/2023

Come for the automation. Stay for the intelligence.



DeepSee™

Table of Contents

Introduction

This white paper offers an overview of AI Model Governance, with a specific focus on how DeepSee uses AI Model Governance to ensure security, transparency, efficacy and efficiency across its AI solutions.

Objective

AI integration demands added scrutiny to support compliance and audits in regulated industries, necessitating tangible evidence of models for informed risk management. Automating regulated workflows and processes requires ingesting documents, messages, images, or even telemetry, and the resulting predictions directly impact outcomes. These outcomes affect more than the bottom line, they affect livelihoods, and therefore necessitate comprehensive AI Governance.

AI governance must be implemented through a framework that defines responsible development, deployment, and use of machine learning (ML) and natural language processing/understanding (NLP/U) models in the context of enhancing operational efficiency. A high-level view of this framework defines governance, risk, and compliance (GRC) for AI implements in order to:

1. **Ensure compliance** with relevant regulations and industry standards across diverse sectors, including data privacy, security, and anti-discrimination laws.

2. **Promote transparency**, fairness, and accountability in the design, development, and deployment of AI systems, enabling stakeholders to better understand and trust the models' decision-making processes.

3. **Safeguard the privacy and confidentiality** of client and company data, while maintaining the highest standards of data quality, integrity, and provenance.

4. **Minimize potential risks and harms** associated with AI systems, including but not limited to model bias, lack of explainability, and unintended consequences by applying best practices and ethical guidelines for AI.

5. **Foster a culture of continuous improvement**, innovation, and collaboration within the organization, driving the adoption of best practices and state-of-the-art techniques in AI model development and governance.

By adhering to these AI Model Governance tenets, DeepSee is committed to delivering reliable, transparent, robust, and ethically sound AI solutions that empower organizations in highly regulated industries to optimize their operational processes and realize greater efficiency, and reliable outcomes.

AI Model Governance

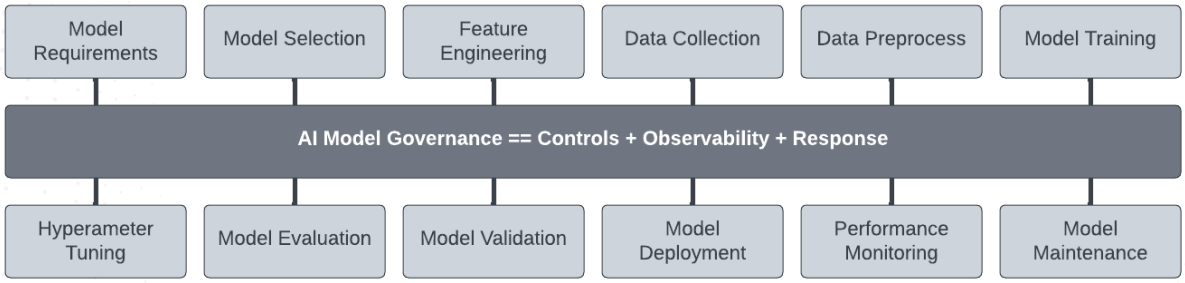
DeepSee observes four key components of AI governance to ensure these tenets are followed. These are: lifecycle management, data governance, audit and review, and continuous learning.

Lifecycle Management

Teams must collaborate effectively and adhere to guidelines and best practices outlined in the AI Model Governance Policy. Clear communication, knowledge sharing, and mutual support are essential for the successful development, automation, deployment, and maintenance of AI models.

The AI Model Lifecycle Management process ensures that AI models are developed, validated, and monitored in a consistent and transparent manner, resulting in improved performance, greater consumer satisfaction, and increased trust in the results/outcomes.

The following components define the key functions in Model Lifecycle Management process:



Guidelines must be established for model validation and performance evaluation, focusing on the selection of evaluation metrics, validation techniques, and benchmark datasets.

Additionally, procedures are specified for model retraining and updating. The approach to model validation and evaluation includes:

- **Evaluation Metrics:** Select appropriate evaluation metrics, such as accuracy, precision, recall, F1-score, or mean squared error, depending on the problem domain and model type.
- **Validation Techniques:** Utilize a variety of validation techniques to ensure robust model evaluation and avoid overfitting. These techniques involve splitting the

dataset into different combinations of training and validation sets to assess the model's generalization capabilities.

- **Benchmark Datasets:** Compare the model's performance against established benchmarks or industry standards where applicable, to ensure that it meets or exceeds expectations.
- **Extensive Automated Testing:** Implement comprehensive automated testing to validate model performance, covering various scenarios and edge cases to minimize the risk of unforeseen issues.
- **Model Retraining and Updating:** Continuously monitor the model's performance in production, retrain and update it as needed, based on new data, customer feedback, or changes in business requirements.

Data Governance

Policies and procedures for data management are defined to ensure the responsible handling of data and compliance with applicable laws, regulations, and industry standards. Data governance practices should encompass data quality, data lineage, data privacy, and data security.

Key aspects of our data governance policies include:

- **Data Collection:** Obtain explicit consent from data subjects for data collection, use, and sharing, in compliance with GDPR, CCPA, and other applicable regulations. This ensures that data is collected responsibly and ethically.

- **Data Storage:** Data storage and processing practices follow industry best practices and relevant ISO/IEC standards. We prioritize the secure storage and handling of data to protect it from unauthorized access, corruption, or loss/leakage.

- **Data Privacy:** We safeguard the privacy of individuals and organizations by implementing privacy-preserving techniques, such as anonymization, data minimization, and differential privacy. These measures help maintain a high level of data privacy while still enabling valuable insights.

- **Data Access:** Access to data is granted on a need-to-know basis, ensuring that sensitive information is only available to those with a legitimate purpose. We maintain logs of data access and usage to track and monitor data handling within our organization.

Through adherence to these data governance practices, you protect the interests of your customers, stakeholders, and the broader community while delivering high-quality machine learning solutions.

Audit and Review

Policies and procedures for data management are defined to ensure the responsible handling of data and compliance with applicable laws, regulations, and industry standards. Data governance practices should encompass data quality, data lineage, data privacy, and data security.

To maintain the highest standards of model governance and ensure compliance with this framework, a periodic review and audit process is implemented. This process enables

evaluation of the effectiveness of governance practices, identifies any potential issues or non-compliance, and uncovers opportunities for improvement.

The audit and review process includes the following steps:

1. **Schedule:** Implement an audit and review process to ensure continuous compliance within the model governance framework and identify opportunities for improvement.

2. **Scope:** Define the scope of each audit and review, including the specific models, processes, and governance elements to be evaluated.

3. **Internal and external audits:** Conduct both internal and external audits to obtain a comprehensive and unbiased assessment of our model governance practices.

4. **Compliance checks:** Verify compliance with applicable laws, regulations, industry standards, and internal policies, identifying any areas where improvements or corrective actions may be necessary.

5. **Performance evaluation:** Assess the performance of our ML models and the effectiveness of our monitoring and maintenance procedures, looking for opportunities to enhance model performance and reliability.

6. **Stakeholder feedback:** Gather feedback from stakeholders, including customers, employees, and partners, to identify any concerns, unmet needs, or suggestions for improvement.

7. **Report findings:** Document and communicate the findings of each audit and review, highlighting any areas of non-compliance, potential risks, or opportunities for improvement.

8. **Action plans:** Develop and implement action plans to address the findings of the audit and review process, ensuring that improvements are made, and any issues are promptly resolved.

9. **Continuous improvement:** Incorporate the insights gained from the audit and review process into our model governance framework and practices, fostering a culture of continuous improvement and adaptation.

Continuous Learning

Emerging technologies, techniques, and best practices develop and progress at an exponential rate. Creating a company culture that embraces and prioritizes continuous learning is the only way organizations can stay ahead of the curve. At DeepSee, we are dedicated to seeking and sharing knowledge across every department. so that we are always up to date with the latest ML technologies and compliance requirements.

This allows us to stay informed about the latest developments in model governance, best practices, ethical considerations, and regulatory compliance. By promoting the sharing of knowledge and experiences through internal discussions, informal presentations, and collaborative problem-solving sessions, we ensure that our teams and technologies remain on the cutting edge.

Conclusion

In conclusion, AI governance is a critical component of any organization that is using or planning to use AI. It encompasses the processes and procedures for ensuring that AI is developed, deployed, and used in a responsible and ethical manner.

There are four key components of AI governance: lifecycle management, data governance, audit and review, and continuous learning.

- Lifecycle management is the process of overseeing the entire life cycle of an AI system, from development to retirement. This includes ensuring that the system is developed and used in accordance with the organization's policies and procedures, and that it is regularly monitored and updated to ensure that it continues to meet the organization's needs.

- Data governance is the process of ensuring that the data used to train and operate AI systems is accurate, complete, and secure. This includes ensuring that the data is collected and stored in a secure manner, and that it is only used for the purposes for which it was collected.

- Audit and review is the process of periodically reviewing AI systems to ensure that they are operating as intended and that they are not being used for malicious purposes. This includes reviewing the system's code, data, and outputs, and interviewing the people who use the system.

- Staying informed through continuous learning and training is the process of ensuring that the people who develop, deploy, and use AI systems are aware of the latest best practices and regulations. This includes attending conferences and workshops, reading industry publications, and taking online courses.

By following these four key components, organizations can help to ensure that AI is used in a responsible and ethical manner.