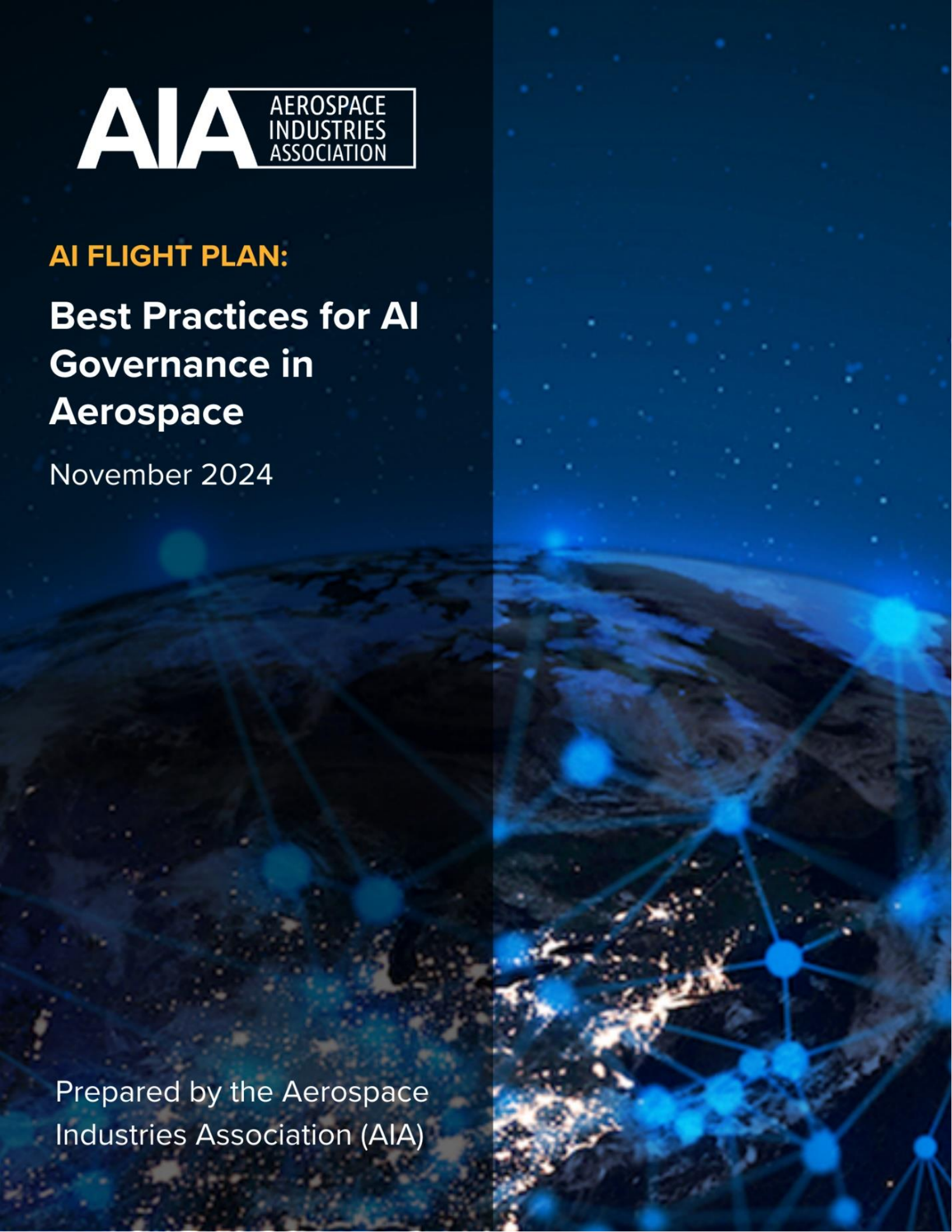
AIA AEROSPACE INDUSTRIES ASSOCIATION

AI FLIGHT PLAN:

Best Practices for AI Governance in Aerospace

November 2024



Prepared by the Aerospace Industries Association (AIA)

Executive Summary

Advanced technology has been the cornerstone of the aerospace industry, and adoption of Artificial Intelligence (AI) is no exception. As AI adoption in aerospace continues to grow, its applications are expanding across the entire product lifecycle and into various business operations, promising unprecedented innovations and efficiencies.

The aerospace industry, characterized by its high-consequence, high-performance nature, demands nothing short of excellence. Here, the stakes are sky-high, and the integration of AI must uphold the rigorous standards that are the hallmark aerospace engineering and operations. Now, more than ever, it is vital to have a discussion around the governing principles of AI.

This paper unveils a comprehensive governance framework, meticulously designed to embed AI into aerospace solutions while exceeding stakeholder expectations.

Crafted by experts from across the Aerospace Industries Association's trail-blazing membership, this paper is not just a guide; it's a flight plan to the future of aerospace. It outlines the elements of an AI governance program tailored specifically for this high-stakes industry, targeting the three primary usages of AI in aerospace: enhancing business operations, aiding in systems design and development processes, and integration into delivered products, software, and services. The anticipated outcomes? Increased stakeholder trust, enhanced transparency, clear regulatory compliance, ethical alignment, superior decision-making, and accelerated decision velocity. Moreover, it provides a robust framework for risk management, ensuring both risk identification and mitigation are top-notch.

By leveraging these governance principles, the aerospace industry can harness the full potential of AI technologies while maintaining its unwavering commitment to safety, quality, reliability, and performance. This balanced approach will pave the way for responsible innovation, allowing the industry to leverage AI in increasingly critical applications as trust and capabilities grow over time.

Table of Contents

Contents

Introduction

Artificial Intelligence (AI) has garnered significant attention in recent years, capturing headlines and sparking debates across various industries. The new interest and reporting on AI could lead one to believe that AI is a new technological development. While there are newly developing AI capabilities, AI has been used within the aerospace sector for many years. Both civil and defense branches of the aviation industry have long been leveraging AI technologies to enhance their operations and capabilities.

The definition of AI varies depending on the source. The Federal Aviation Administration (FAA) describes AI as "a discipline of creating behavior that mimics aspects of human decisions at the machine level." This definition emphasizes AI's ability to replicate human-like decision-making processes. On the other hand, the Organisation for Economic Co-operation and Development (OECD) offers a more comprehensive definition: "An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment." This definition highlights the input-output nature of AI systems and their potential impact on both physical and virtual realms.

The aerospace industry faces unique challenges when it comes to implementing AI technologies. Because the safety of passengers and crew is paramount, the stakes are exceptionally high within commercial aviation. One of the primary obstacles is the difficulty in explaining the reasons behind the AI model outputs, which is crucial in an industry where transparency and accountability are essential. Other challenges include the availability, privacy, and intellectual property (IP) rights for the data used to train AI models; security in collaborating on AI-based design with third parties; and safety mechanisms for monitoring and controlling AI functions, particularly in response to unforeseen stimulus not included in their training data.

To address these challenges and ensure the responsible use of AI in aerospace, the concept of AI Governance has emerged as a critical enabler. AI governance sets organizational expectations that provide the greatest chance for achieving reliable and repeatable processes. It encompasses a set of policies, procedures, and practices that an organization establishes to ensure legal, ethical, and aligned outcomes in an AI implementation.

Governance plays a vital role in outlining a structure for decision-making and accountability across an organization. It is an integral part of enterprise risk management, helping to mitigate potential issues and ensure consistency in AI application. The expectations and guidelines discussed in this white paper apply not only to AI used in products, services, and software intended for customer use, but also to AI systems designed for internal organizational purposes.

As we delve deeper into the role of AI in the U.S. aviation industry, it's essential to keep these foundational concepts in mind. The following paper will explore specific applications, challenges, and best practices for AI implementation in aerospace, always with an eye toward responsible innovation and unwavering commitment to safety.

Governance Program Elements

Policy

The cornerstone of successful AI implementation in the aerospace industry lies in effective AI governance, which is fundamentally rooted in the development and active application of comprehensive policies. To establish a robust AI governance framework, it is crucial to begin with a clear foundation that outlines program goals, identifies potential risks, and recognizes key stakeholders.

For aerospace solution providers, we strongly recommend adopting a tailored, enterprise-wide policy centered on the responsible use of AI. In this context, "responsible" encompasses a multifaceted approach that ensures AI is designed, developed, deployed, and utilized in a manner that is safe, secure, ethical, legal, and socially responsible, with policies consistently applied across the organization. This holistic view of responsibility helps to address the unique challenges and high stakes inherent in the aerospace industry.

The policy should be built upon clear principles for "Responsible AI." These principles serve as guiding pillars for all AI-related activities within the organization. We propose the following key principles:

- Trust: This principle emphasizes the importance of designing, developing, and operating AI systems that can be trusted by all stakeholders, including customers, suppliers, employees, regulators, investors, and society at large. Trust is paramount in the aerospace industry, where safety and reliability are non-negotiable.
- Accountability: This principle underscores the necessity of designing, developing, operating, and deploying AI systems in full compliance with applicable legal and regulatory frameworks. It ensures that the organization takes responsibility for its AI systems and their impacts.
- Objectivity: To uphold this principle, organizations must take deliberate steps to minimize unintended bias in their AI systems. This is crucial in ensuring fair and equitable outcomes, especially in critical decision-making processes.
- Transparency and Traceability: This principle advocates for the use of explainable and auditable methods in AI development and deployment. In an industry where the reasoning behind decisions can be a matter of life and death, the ability to trace, explain, and audit AI outputs is essential. Transparency requires instrumentation in product design and operations, so system data can be recorded, shared, and evaluated to continually optimize safety and performance.
- Safety and Reliability: This principle emphasizes the importance of using AI capabilities within explicit, well-defined use cases. It also calls for adopting a systemic risk-management approach to ensure that systems perform as intended. This is particularly critical in aerospace applications where system failures can have catastrophic consequences.

- Security: This principle emphasizes the importance of security in the entire AI life cycle from model development and deployment to continuous improvement. This is one of the important considerations for aerospace industry.
- Governance: This principle ensures oversight and accountability at all levels of the organization – corporate, functional, and business. It establishes a clear chain of responsibility and decision-making processes for AI-related matters.

By grounding AI governance policy in these principles, aerospace organizations can create a framework that not only guides the responsible development and use of AI, but also aligns with the industry's stringent safety and reliability requirements. This approach helps to balance innovation with risk management, ensuring that AI technologies are leveraged to their full potential while maintaining the highest standards of safety and ethical consideration.

As the aerospace industry continues to evolve and integrate more AI-driven solutions, having a well-defined AI governance policy will be crucial in navigating the complex landscape of technological advancement, regulatory compliance, and public trust. It provides a roadmap for responsible innovation, helping organizations to harness the power of AI while upholding their commitment to safety, security, ethics, and social responsibility.

AI use case catalog

In the rapidly evolving landscape of AI in aerospace, it is crucial for organizations to maintain a comprehensive and up-to-date catalog of AI use cases. This catalog should serve as a central repository of knowledge, documenting the various ways AI is being implemented across the organization. The creation and maintenance of such a catalog is not just an administrative task, but a strategic initiative that can significantly enhance an organization's AI governance and innovation capabilities.

The primary purpose of an AI use case catalog is to effectively manage the portfolio of AI implementations within an organization. By systematically documenting each use case, including its objectives, methodologies, outcomes, and lessons learned, the organization creates a valuable resource that can inform decision-making, guide future developments, and ensure consistency in AI application across different departments or projects.

Such a catalog offers multiple benefits that extend far beyond simple documentation. First, it can serve as a crucial tool for future risk assessments. By having a detailed record of past and current AI implementations, organizations can more easily identify potential risks and challenges associated with new AI projects. This proactive approach to risk management is particularly important in the aerospace industry, where safety is paramount. It provides a solid foundation for AI program management by tracking the progress and impact of AI initiatives over time through data-driven evaluation, with a comprehensive view of the organization’s overall AI capabilities and gaps. As organizations continually review the catalog, it can be a source of inspiration for new innovations.

Secondly, an AI use case catalog provides a solid foundation for AI program measurement. It allows organizations to track the progress and impact of their AI initiatives over time, enabling them to assess the effectiveness of different approaches and technologies. This data-driven evaluation can

help in optimizing resource allocation, identifying areas for improvement, and demonstrating the value of AI investments to stakeholders.

Furthermore, the catalog can inspire new innovations. By capturing the details of successful (and even unsuccessful) AI implementations, it provides a rich source of ideas that can be adapted, combined, or expanded upon to create novel solutions. This cross-pollination of ideas can accelerate innovation and help the organization stay at the forefront of AI application in aerospace.

Beyond these core benefits, an AI use case catalog can also:

- Facilitate knowledge sharing across the organization, breaking down silos and promoting collaboration.
- Support training and onboarding of new team members by providing real-world examples of AI applications.
- Aid in regulatory compliance by maintaining a clear record of AI systems in use and their purposes.
- Inform strategic planning by providing a comprehensive view of the organization's AI capabilities and gaps.
- Enhance communication with external stakeholders, including customers, partners, and regulators, by offering concrete examples of the organization's AI expertise and responsible implementation practices.

To maximize the value of an AI use case catalog, it should be regularly updated and easily accessible to relevant stakeholders. It should include not only technical details but also business outcomes, ethical considerations, and lessons learned. Organizations might consider implementing a standardized template for documenting use cases to ensure consistency and completeness of information. A representative list of aspects that should be covered in the catalog is documented below:

- Use case name
- Business problem statement
- Use case justification
- Business benefits envisaged and realized
- Applicable business unit and/or business process
- Intended function, criticality, and planned functional domain
- Data sources and data strategy
- Model selection, model development and training, and model management approach
- Responsible AI considerations
- Lifecycle management
- Lessons learned

An AI use case catalog is a powerful tool that goes beyond mere documentation. It serves as a strategic asset that supports risk management, performance measurement, innovation, and organizational learning. For aerospace companies navigating the complex intersection of AI and aviation, such a catalog is an invaluable resource in ensuring the safe, effective, and responsible implementation of AI technologies.

Risk management

Enterprise Risk Management (ERM) is a comprehensive framework of coordinated activities designed to guide an organization's approach to risk across all its operations. This broad discipline encompasses various specialized areas such as supply chain risk management, financial risk management, human resource risk management, and business conditions risk management. Within this expansive framework, AI risk management has emerged as a crucial component, particularly for organizations in the aerospace industry where the stakes are exceptionally high. Functional risk management is a subset of ERM, and each functional area has an obligation to manage specific risks accordingly.

AI risk management is a systematic and comprehensive approach to identifying, assessing, managing, and monitoring risks associated with the use of AI throughout an organization. This includes AI applications in developing new products, software, or services; AI used in back-office operations; and AI integrated into customer-delivered products, software, or services. The scope of AI risk management is intentionally broad to capture all potential risk areas associated with AI implementation.

A key aspect of AI risk management is the establishment of AI governance. This involves documenting an organization's ethical standards and policies for AI use, ensuring that AI systems are designed and operated in alignment with shared values and legal requirements. In the aerospace industry, where safety is paramount and regulatory compliance is stringent, this governance is crucial in preventing ethical breaches, safety issues, security vulnerabilities, and legal complications. These potential problems represent significant risks that could have far-reaching consequences if not properly managed.

To effectively manage AI risks, enterprises should establish a framework to classify the various types of risks associated with AI use within the organization. This classification system helps in prioritizing risks and allocating resources for their mitigation. It's important to note that risk management is not a one-time activity but an ongoing process. There should be an expectation and a mechanism in place for reviewing risks as new information emerges or as the AI landscape evolves, with clear ownership assigned to risks and mitigation efforts.

A critical component of effective AI risk management is the identification and involvement of key stakeholders. These may include AI developers, system users, key information technology leaders, legal experts, ethics officers, and representatives from various business units. Organizations should establish a regular process by which these stakeholders meet to review risks, track the progress of mitigation efforts, and make decisions on risk mitigation strategies. This collaborative approach ensures a comprehensive view of AI risks and promotes the consistent reduction of risks or the development of alternative mitigation strategies.

Lastly, it's crucial to recognize that risk management extends beyond prevention and mitigation to include planning for disaster response and recovery. In the context of AI, this might involve scenarios such as AI system failures, data breaches, or unintended consequences of AI decisions. Having well-defined protocols for responding to and recovering from such incidents is an essential part of a robust AI risk management strategy. In addition, where feasible, proactively implementing human-in-loop mechanisms for high-risk AI applications can help reduce disaster risk.

Training data evaluation

In the realm of AI development, the evaluation and management of training data is crucial. The content of the training data essentially becomes the system specification, as it fundamentally determines the outcomes of the AI system. Consequently, the process of data selection is tantamount to defining system requirements, underscoring the critical nature of this phase in AI development.

This process is not without its pitfalls. Developers must be vigilant about potential issues such as data bias, incompleteness, or inconsistencies that could lead to skewed or unreliable AI outputs. Best practices in training data evaluation include thorough data audits, diversity checks, and ongoing monitoring of data quality and relevance in order to identify data drift and its implications. Selection of training data should include consideration of the intended operating design domain (ODD) of the model. By including considerations of the ODD in the training data set, the model output or behavior can be limited to acceptable operating limits.

Data governance takes on heightened significance in regulated environments like aerospace, serving as a key differentiator from less stringent sectors. Robust data governance frameworks ensure data integrity, traceability, and compliance with industry standards and regulations. This is particularly crucial when considering certification requirements for AI in aerospace applications, such as meeting the stringent standards set forth in DO-178C "Software Considerations in Airborne Systems and Equipment Certification," for software considerations in airborne systems and equipment certification.

Understanding and addressing bias in training data is crucial. This involves not only identifying existing biases but also anticipating potential biases that could emerge. Regular reevaluation of data sets is necessary as new expectations, regulations, and ethical considerations evolve over time.

Ensuring data is thorough and covers all potential situations is critical for enhancing AI model resilience. However, some edge cases may lack data, which necessitates the use of physics-based models to create synthetic data and mitigate these data gaps.

Privacy and confidentiality are also major concerns in training data evaluation. Processes must be established to protect personal and confidential information, including methods for anonymization or elimination of sensitive data before its use in training. This ties into broader enterprise data architecture discussions, raising questions about how organizations can generate, store, secure, and catalog data for AI training at an organizational level. Decisions about data ownership, particularly for critical data sets, need to be clearly defined.

Intellectual property considerations add another layer of complexity to training data evaluation. Organizations must clearly delineate which data sources are permissible for use, and which are forbidden. Furthermore, policies for sharing data with third parties need to be established to protect proprietary information while enabling necessary collaborations.

To ensure continuous improvement of AI systems, it's advisable to establish a program for gathering data from as many relevant sources as practical on an ongoing basis. This approach allows for the refinement and updating of AI models to maintain their accuracy and relevance over time. Training

data must be stored through the entire lifecycle of the product, like software source code, so that as product design changes are made, or if failure analysis is needed, the organization can recreate, evaluate, and update the AI-based function. The governance policy should address how this data is stored, configuration-managed, and secured, including provisions for disaster recovery.

Model assessment, selection, or development process

The process of model selection is a critical task in developing an AI solution, particularly in the aerospace industry where predictability is paramount. The spectrum of available models is vast, ranging from simple regressions with limited inference capabilities, to sophisticated multi-modal large language models. Each type of model comes with its own set of strengths, limitations, and potential risks.

To ensure consistency and rigor in this crucial decision-making process, AI program governance should provide clear directions on the model selection process. These guidelines should offer insights into the risks and necessary controls associated with different model types or techniques. However, it's important that these directions maintain appropriate flexibility, avoiding rigid decision logic that may not be applicable across all applications or areas within the organization. Model selection can be heavily influenced by constraints such as cost, licensing requirements, sustainability, computational resources, latency, data privacy concerns, data availability, potential data bias, storage costs, ethical considerations, model explainability requirements, scalability needs, and intellectual property exposure risks.

The model selection process should begin with a clear definition of the problem to be solved or the objectives of the project. This foundational step ensures that the chosen model aligns with the specific needs of the task at hand. Understanding the anticipated outcome is equally crucial in guiding the selection process. For instance, the project may require prediction, classification, clustering, or natural language processing capabilities, each of which may be better served by different types of models.

A key aspect of model selection is defining the assessment criteria - the metrics and standards by which the model's performance will be judged. These criteria help determine if a model is "good enough" for the intended application. Assessment criteria can be both subjective and objective. Subjective criteria might include factors like explainability and transparency, which are particularly important in aerospace applications where the reasoning behind AI decisions may need to be audited or explained to regulators or end-users. Objective criteria, on the other hand, include quantifiable metrics such as accuracy, precision, and recall.

It is essential to include a plan for testing bias in the resulting model as part of the assessment process. This helps ensure that the model doesn't perpetuate or amplify existing biases, which could lead to unfair or potentially dangerous outcomes in aerospace applications.

Lastly, just as with evaluating training data, considerations for future system certification should be incorporated into the model selection process. This forward-looking approach ensures that the chosen model not only meets current needs, but also aligns with potential future regulatory requirements or certification standards in the aerospace industry. Models and associated “DevOps” frameworks used in product design must be stored through the entire lifecycle of the

product, like software source code. Then if product design changes are made, or if failure analysis is needed, the organization can recreate, evaluate and update the AI-based function. The governance policy should address how models are stored, configuration-managed, and secured, including provisions for disaster recovery.

In cases where bespoke models are developed to meet specific operational or validation purposes, all considerations above will also apply. Users should work through the same expectations of problem definition, assessment criteria, testing for bias, and considerations for future certification requirements as the model is built.

Test, verification, and certification of AI in aerospace systems

The aerospace industry has long been characterized by its stringent safety standards supported by rigorous certification processes. As AI increasingly becomes an integral part of aerospace systems, the industry faces new challenges in testing, verifying, and certifying these AI-enhanced technologies. This is particularly crucial in high-consequence environments where system failures could lead to catastrophic outcomes.

Traditionally, aerospace systems are tested against clearly defined specifications, with testing protocols designed to support established certification processes. These processes have been refined over decades to ensure the utmost safety and reliability of aircraft and related systems. However, the introduction of AI into these systems necessitates a reevaluation and adaptation of these long-standing practices.

Planning for the testing, verification, and certification of AI-enhanced aerospace systems must consider several unique aspects of AI:

- Non-deterministic behavior: Unlike traditional software systems, AI systems can produce different outputs for the same input under certain conditions. This variability needs to be accounted for in testing protocols.
- Explainability, traceability, and transparency: Certification processes may require a clear understanding of how AI systems arrive at their decisions. This can be challenging with complex AI models like deep neural networks.
- Continuous learning systems: Some AI systems are designed to learn and adapt over time. Testing and certification processes need to account for how these systems might evolve in operation.
- Edge cases and unexpected scenarios: AI systems need to be rigorously tested for their response to rare or unforeseen situations, which is particularly critical in aerospace applications.
- Data dependency: As noted earlier in this document, AI system performance is heavily influenced by their training data. Certification processes may need to include evaluation of training data quality and representativeness.
- Robustness and adversarial attacks: AI systems must be tested for their resilience against potential manipulations or adversarial inputs that could compromise their performance.

To address these challenges, several approaches can be considered:

- Simulation-based testing: Extensive use of simulations to test AI systems across a wide range of scenarios, including edge cases.
- Formal verification methods: Application of mathematical techniques to prove certain properties or behaviors of AI systems.
- Scenario-based testing: Development of comprehensive test scenarios that cover both expected and unexpected operational conditions.
- Monitoring and logging: Implementation of robust monitoring systems to track AI performance and decision-making in real time during operations.
- Hybrid approaches: Combining AI with traditional rule-based systems to leverage the strengths of both while mitigating risks.
- Ethical and bias testing: Incorporating tests to ensure AI systems make fair and unbiased decisions, particularly in safety-critical situations.
- Security testing: Test plans should include assessments of security risks and associated tests to ensure mitigations are in place.
- Incremental certification: Developing processes for continuous or phased certification as AI systems learn and evolve.
- Collaborative industry standards: Working with regulatory bodies and industry partners to develop new standards and best practices specifically for AI in aerospace applications.

Moreover, organizations should consider establishing dedicated AI assurance teams that bring together expertise in both AI and traditional aerospace certification processes. These teams can bridge the gap between cutting-edge AI development and the rigorous safety requirements of the aerospace industry.

Ethical standards

Ethical considerations form a cornerstone of any robust AI governance program, particularly in the aerospace industry where decisions made by AI systems can have profound implications for human safety and well-being.

A comprehensive AI governance program should include use case tailored assessments of fairness and bias. This involves scrutinizing AI models and their outputs to ensure they do not discriminate against or unfairly impact any particular group or individual. For instance, in the context of aerospace, this could mean ensuring that AI-driven security screening systems at airports do not exhibit bias based on race, gender, or nationality. Equally important are the development and implementation of approaches to mitigate identified biases. This might involve retraining models with more diverse datasets, implementing fairness constraints in algorithms, or employing ensemble methods that combine multiple models to reduce individual biases.

Data privacy and security are integral components of ethical standards in AI governance. In the aerospace sector, where sensitive information abounds – from passenger data to proprietary aircraft designs – robust data protection measures are non-negotiable. Privacy concerns that should be addressed include:

- Data minimization: In cases where Personally Identifiable Information (PII) is available, collecting and retaining only the data necessary for the specific AI application.

- Anonymization: Removing or encrypting PII to protect individual privacy.
- Informed consent: Ensuring that individuals are aware of how their data is being used and have given permission for its use.
- Data security: Implementing strong cybersecurity measures to protect against unauthorized access or data breaches.

Beyond these considerations, ethical standards in AI governance for aerospace should also address:

- Transparency and explainability: Ensuring that AI decision-making processes can be understood and audited, especially in safety-critical applications.
- Accountability: Establishing clear lines of responsibility for AI system outcomes.
- Human oversight: Maintaining appropriate human control and intervention capabilities in AI systems.
- Safety and reliability: Prioritizing the safety of passengers, crew, and the public in all AI applications.
- Environmental impact: Considering the ecological footprint of AI systems and striving for sustainability.
- Ethical use of AI in defense applications: Establishing clear guidelines for the use of AI in military aerospace technologies.

An essential outcome of establishing these ethical standards is the development of standardized guardrails for Artificial Intelligence/Machine Learning (AI/ML) systems. These guardrails serve as built-in safeguards that ensure AI systems operate within predefined ethical boundaries. They might include:

- Algorithmic checks to prevent biased outputs.
- Automatic system shutdowns if ethical thresholds are breached.
- Mandatory human approval for critical decisions.
- Regular ethical audits and impact assessments.
- Continuous monitoring for unintended consequences or ethical drift.

Human oversight expectations

Human oversight and control have been longstanding pillars of aviation safety and operation. As artificial intelligence increasingly permeates aerospace systems, the industry faces the challenge of integrating these new technologies while maintaining appropriate human involvement. We view this change as a significant step forward for the industry – the increased automation from AI should increase reliability and capability of systems, while freeing the human workforce to perform higher level, more value-added tasks. Balance between automation and human oversight is crucial for ensuring safety, accountability, and public trust in AI-enhanced aerospace applications.

Traditionally, the aerospace industry has relied on deterministic models, which offer the benefit of explainability. Engineers and operators can trace the logic in order to understand why a system made a particular decision. However, with the advent of probabilistic models and neural networks,

which form the basis of many advanced AI systems, explanation becomes much more complex and, in some cases, may be impossible. This shift necessitates a reevaluation of how human oversight is implemented and maintained.

Defining clear expectations for human oversight is crucial for the success of any AI implementation in aerospace. These expectations should be tailored to the specific application and potential risks involved. Organizations must carefully consider the appropriate level of human involvement for each AI system, weighing the benefits of automation against the need for human judgment and intervention.

In this context, it's essential to distinguish between different modes of human involvement:

- Human-in-the-loop: This approach involves active human participation in the decision-making process. The AI system may provide recommendations or perform certain tasks, but a human operator makes the final decisions or takes critical actions.
- Human-on-the-loop: In this mode, the AI system operates autonomously, but under human supervision. The human operator monitors the system's performance and can intervene if necessary.
- Fully autonomous operation: Here, the AI system operates independently without direct human oversight. This mode is typically reserved for low-risk, well-defined tasks, or environments where real-time human intervention is not feasible.

Each AI implementation in aerospace should carefully consider which of these approaches is most appropriate for different applications or aspects of the system. For instance, an AI system managing in-flight entertainment might operate autonomously, while one assisting with navigation or flight control would likely require a human-in-the-loop approach.

AI implementations should make explicit choices about goals and decision-making processes for both humans and machines. Attention must be given to transitions of control between machine and human. These transition points are critical, due to the potential of failure at the point of handoff. The system must ensure that when control shifts from AI to human or vice versa, each entity is capable of accommodating the new tasking. This involves considerations such as situational awareness, response time, and cognitive load.

The importance of human oversight in AI systems is underscored by directives like the National Security Memorandum, which emphasizes the safe and ethical use of AI in military and intelligence operations. This includes ensuring that AI can be safely and effectively controlled to prevent unintended or harmful actions. For example, the concept of an "override/kill switch" for autonomous weapons systems highlights the critical nature of maintaining ultimate human control over AI in high-stakes scenarios.

As part of implementing human oversight, organizations should also identify specific tasks or job functions that they will not allow AI to perform. This delineation helps maintain clear boundaries and ensures that critical decisions remain in human hands. These restrictions might include final go/no-go decisions for launches, override controls, emergency response coordination, or ethical judgments in unforeseen scenarios.

AI program measurement

In the rapidly evolving field of AI, establishing robust measurement and monitoring protocols is critical for ensuring the ongoing effectiveness, safety, and reliability of AI implementations. A comprehensive AI program measurement strategy should encompass not only individual AI implementations but also the overarching AI governance program itself.

To achieve this, organizations should establish clear processes, metrics, and timelines for auditing and monitoring AI implementations. These should be designed to assess various aspects of AI performance, including accuracy, efficiency, safety, and adherence to ethical standards. The frequency of these audits may vary depending on the criticality of the AI system, with safety-critical applications potentially requiring more frequent assessments.

Metrics for individual AI implementations might include:

- Accuracy rates in decision-making or predictions
- Response times and system latency
- Frequency and nature of errors or unexpected outputs
- Bias detection and fairness measures
- Resource utilization and efficiency

For the overall AI program, metrics could focus on:

- Number of successful AI implementations
- Return on investment for AI projects
- Compliance with regulatory requirements
- Progress in AI-related skills development within the organization
- Effectiveness of AI governance policies

User feedback should be an integral part of both program and implementation measurement. This feedback provides valuable insights into the real-world performance and usability of AI systems. It can highlight issues that may not be apparent through quantitative metrics alone, such as user trust in the system or ease of interaction. Regular surveys, interviews, or feedback sessions with end-users can provide this qualitative data, which should be systematically collected, analyzed, and acted upon. Where possible and appropriate, system data should be logged and correlated to user feedback to support creation of new training data and system performance improvements.

An important aspect of AI program measurement in aerospace applications is ensuring that the AI systems operate within the bounds of their training data and that their behavior conforms to expectations. This requires specific guidance and rigorous system and model monitoring protocols. Organizations should implement continuous monitoring systems that can:

- Track the input data to ensure it falls within the range of the training data set.
- Monitor system outputs for anomalies or unexpected behaviors.
- Detect concept drift, where the relationship between input and output data changes over time.

- Alert human operators when the system encounters scenarios outside its training or when its confidence in decisions falls below a predetermined threshold.

This monitoring should be coupled with clear procedures for human intervention when necessary. For instance, if an AI system encounters a scenario significantly different from its training data, it should either defer to human judgment or follow pre-defined safety protocols.

Additionally, organizations should establish a process for regular revalidation of AI models. This is particularly important in the dynamic aerospace environment, where new technologies, regulations, or operational scenarios may emerge that were not considered in the original training data.

Legal compliance expectations and processes

The legal and regulatory landscape surrounding artificial intelligence in aerospace is evolving and will likely continue to do so for the foreseeable future. This reality necessitates a proactive and adaptable approach to legal compliance in AI governance programs. This dynamic environment presents both challenges and opportunities for aerospace companies implementing AI solutions.

To navigate this complex and shifting terrain, collaboration with legal experts should be an intentional and integral part of any AI governance program. This collaboration should not be a one-time consultation but an ongoing partnership. Legal experts can provide invaluable insights into current regulations in relevant geographies, help interpret new laws as they emerge, and assist in forecasting potential future legal requirements. They can also help organizations structure their AI development and deployment processes in ways that facilitate compliance and reduce legal risks.

Moreover, policy in this area continues to evolve, and it is incumbent upon stakeholders such as aerospace companies to actively engage with regulatory bodies. This engagement should be a committed part of their AI compliance program.

Furthermore, companies should establish internal processes to regularly review and update their AI governance policies considering new legal developments. This might involve creating a dedicated team or task force responsible for monitoring legal changes, assessing their impact on the company's AI initiatives, and implementing necessary adjustments to ensure ongoing compliance.

AI program security

AI program security is a critical component of the broader enterprise security framework in aerospace organizations. It's essential to establish strong alignment between AI-specific security measures and the overall enterprise security strategy to ensure comprehensive protection against potential threats.

For aerospace applications, there must be seamless integration between the product or system security plan and the AI security plan. This alignment ensures that AI components are not treated as separate entities but are fully incorporated into the overall security architecture of aerospace systems.

One effective method for detecting and mitigating potential security threats is model explainability. By making AI models more transparent and interpretable, organizations can more easily identify

anomalies or unexpected behaviors that might indicate the presence of bad actors or security breaches.

Security considerations should begin at the data level. Training data used for AI models should be rigorously assessed for potential tampering or poisoning by malicious actors' intent on influencing model outcomes. This is particularly crucial in aerospace applications where compromised data could lead to serious safety risks.

Once models are developed, they must be secured to prevent unauthorized alterations. This includes implementing robust access controls, encryption, and audit trails to track any changes made to the models.

Standard practices in configuration management should be applied to data inputs, resulting models, and development tool chains. This ensures version control, traceability, and the ability to roll back to previous versions if security issues are detected.

Organizations should establish clear plans and protocols for preventing data breaches and handling such events if they occur. This includes defining response teams, communication protocols, and recovery procedures.

Prompt injection attempts can be resisted such as input sanitization, context management, permissions management, and anomaly detection.

It's important to recognize that security threats can target any part of the AI/ML lifecycle and supply chain, including the data acquisition and ML training phases. Therefore, security measures must be implemented across the entire AI development and deployment process.

The significance of AI security in aerospace, particularly in defense applications, is underscored by bodies like the Defense Science Board. They recognize autonomy in adversarial physical and information systems as a critical threat in future military operations, highlighting the need for robust AI security measures.

For guidance on specific security considerations, especially for large language models, resources like the Open Worldwide Application Security Project Top 10 List for Large Language Models provide valuable insights. While not all items on this list may apply directly to every aerospace AI application, it offers a comprehensive overview of potential security vulnerabilities that should be considered.

Roles and responsibilities

Clearly delineating roles and responsibilities within an AI program is crucial for ensuring both accountability and efficiency. A well-structured governance framework for AI initiatives helps to streamline decision-making processes, clarify lines of authority, and ensure that all aspects of AI development and deployment are adequately overseen.

One of the key aspects of this governance structure is the explicit recognition of the natural tension that exists between assurance and velocity in AI development. On one hand, the aerospace industry demands rigorous testing and validation to ensure safety and reliability. On the other hand, there's a need for rapid innovation and deployment to stay competitive and leverage the benefits of

AI technologies. By acknowledging this inherent competition, management and users are better equipped to make informed decisions about AI usage, balancing the need for thoroughness with the imperative for progress and speed.

To manage this balance effectively, organizations should clearly define different roles within the AI program. These might include:

- AI Development Teams: Responsible for designing, developing, and implementing AI solutions.
- AI Ethics Committee: Oversees the ethical implications of AI applications and ensures compliance with ethical guidelines.
- Quality Assurance Teams: Responsible for testing and validating AI systems against safety and performance standards.
- Legal and Compliance Officers including Certification Engineers: Ensure AI initiatives comply with relevant laws and regulations.
- Business Stakeholders: Provide input on business requirements and use cases for AI applications.
- Security Teams: Oversee the cybersecurity aspects of AI systems.
- Data Governance Teams: Manage data quality, privacy, and usage across AI initiatives.

Each of these roles should have clearly defined responsibilities, authority levels, and reporting structures. This clarity helps to prevent duplication of efforts, ensures comprehensive coverage of all aspects of AI governance, and facilitates swift decision-making when necessary.

An essential component of the governance framework is the establishment of risk categories for AI implementations. These categories should be accompanied by defined protocols for stakeholder engagement at different risk levels. For instance, low-risk AI applications might require minimal stakeholder involvement, while high-risk applications might necessitate approval from senior management, ethics committees, and potentially even external regulatory bodies.

Furthermore, each AI implementation should include specific plans for stakeholder engagement and approvals. This ensures that all relevant parties are involved at appropriate stages of the project, from initial concept development through to deployment and ongoing monitoring. These plans should outline:

- Which stakeholders need to be involved at each stage of the AI lifecycle
- The nature and frequency of stakeholder communications
- Decision-making processes and approval requirements
- Escalation procedures for addressing concerns or conflicts

Change management and the workforce

The integration of artificial intelligence into operations represents a significant shift in how work is performed, decisions are made, and systems are managed. This change is not only extensive but also rapidly evolving, presenting unique challenges for workforce adaptation and organizational culture. Recognizing and effectively managing this change is crucial for the successful implementation of AI programs.

One of the most distinctive aspects of AI-driven change is its potential for abruptness. Unlike many technological advancements of the past, which often allowed for gradual adaptation, AI has the capacity to transform roles and processes swiftly and dramatically. This rapidity of change is a unique characteristic that must be explicitly acknowledged and addressed in AI change management strategies. Traditional change management approaches may need to be accelerated and modified to keep pace with the rapid evolution of AI technologies and their impacts on the workforce.

It's natural and understandable for employees to have concerns about how AI will affect their work. These concerns may range from job security fears to uncertainties about new skill requirements or changes in job responsibilities. Some employees may worry about being replaced by AI systems, while others might be anxious about their ability to adapt to new AI-driven processes. Addressing these concerns openly and empathetically is crucial for maintaining morale and fostering a positive attitude towards AI adoption.

Clear and comprehensive communication is paramount in managing this change effectively. Organizations should prioritize transparent and regular communication about:

- The specific changes being implemented: Clearly outline how AI is being integrated into various processes and systems, and how this will affect different roles and departments.
- The benefits of these changes: Articulate how AI implementation will benefit not only the organization but also individual employees. This might include improved efficiency, reduced repetitive tasks, enhanced decision-making capabilities, or new opportunities for skill development and career growth.
- Expectations for employees: Provide clear guidance on what is expected from employees during and after the AI implementation. This might include participation in training programs, collaboration with AI systems, or taking on new responsibilities that leverage AI capabilities.
- Timeline and support: Offer a clear timeline for AI implementation and the associated changes. Importantly, communicate the support mechanisms available to employees, such as training programs, mentoring, or resources for learning about AI.
- Long-term vision: Share the organization's long-term vision for AI integration and how it aligns with the company's overall mission and goals. This can help employees see the bigger picture and understand their role in this evolving landscape.

Sustainability

The integration of AI into sustainability programs should be holistic, taking into account both the positive and negative impacts of AI development and usage. On the positive side, AI has the potential to significantly enhance sustainability efforts in aerospace. For instance, AI can optimize flight routes for fuel efficiency, improve predictive maintenance to reduce waste and extend the lifespan of aircraft components, and enhance the design process for more environmentally friendly aircraft.

However, it's equally important to consider the potential negative impacts of AI on sustainability. The development and operation of AI systems can be energy-intensive, contributing to increased carbon emissions if not managed properly. There's also the consideration of the environmental

impact of the hardware required for AI systems, including the mining of rare earth metals for components and the disposal of electronic waste.

Organizations should conduct thorough assessments of the environmental impact of their AI initiatives, from the energy consumption of data centers used for AI training to the lifecycle analysis of AI-enabled products. This balanced approach ensures that the pursuit of AI innovation doesn't come at the cost of environmental sustainability.

Moreover, the influence of AI on sustainability extends beyond the immediate operations of aerospace companies to encompass the entire supply chain. AI can play a crucial role in optimizing supply chain logistics, reducing waste, and improving energy efficiency across the entire value chain. For example, AI-driven demand forecasting can help reduce overproduction and associated waste, while AI-optimized logistics can minimize transportation-related emissions.

Conclusion

In conclusion, the integration of AI in the aerospace industry is another tool in the long line of high-technology solutions utilized by the aerospace industry. Effective AI governance is crucial to ensure that these technologies are developed and deployed responsibly, with a focus on safety, transparency, and ethical considerations. By establishing robust frameworks and regulatory standards, the aerospace sector can harness the full potential of AI while mitigating risks and addressing public concerns. Continued collaboration between industry, the many aerospace stakeholders, and policymakers will be essential in navigating the complexities of AI governance while fostering innovation in a manner that benefits society as a whole.