Discovery & Inventory

- **Are all GenAI models and systems (in use or development) inventorised?: Yes, fully**

- **Are data sources (training, fine-tuning, RAG) for GenAI models documented and assessed for bias or quality?: Documented, not assessed**

- **Is each GenAI system's purpose and operational context clearly defined and documented?: For most systems**

- **If using third-party GenAI models/APIs (e.g., LLMs), are their compliance and risk profiles understood?: Yes, but not understood**

- **Are system boundaries and integration points for GenAI applications with other enterprise systems defined?: Partially defined**

Regulation & Classification

- **Are your GenAI systems classified by EU AI Act risk categories (unacceptable, high, limited, minimal)?: Classification in progress**

- **If GenAI systems are high-risk, are you prepared for EU AI Act obligations (e.g., QMS, technical documentation, conformity assessment)?: Partially prepared**

- **If developing/using GPAI models, are you aware of specific EU AI Act obligations (e.g., transparency, technical documentation)?: Unaware of specific obligations**

- **For high-risk GenAI systems, is there a plan for conformity assessment before market placement or use?: Planning in progress**

Impact Assessment & Mitigation

- **Has a Fundamental Rights Impact Assessment (FRIA) been conducted for GenAI systems, particularly if high-risk?: FRIA in progress**

- **Are processes in place to detect, document, and mitigate biases in GenAI models and their outputs?: Processes in development**

- **Have potential societal impacts (e.g., employment, public discourse, environment) of GenAI systems been assessed?: Aware, not formally assessed**

- **Has the potential for misuse or malicious use of GenAI systems (e.g., deepfakes, disinformation) been assessed and documented?: Assessed, not documented**

- **Has the environmental impact (e.g., energy use for training/inference) of GenAI models been considered or assessed?: Considered, not formally assessed**

Governance & Operations

- **Has your organization established a formal, documented policy framework specifically addressing the governance of GenAI development, deployment, and use, including acceptable use, data handling, and ethical considerations?: Policy in development**

- **Are there clearly defined roles, responsibilities, and accountability structures for the oversight and governance of GenAI systems, including a designated individual or body responsible for GenAI compliance?: Partially defined or assigned**

- **Is there a formal process to regularly review and ensure that GenAI systems and their use comply with applicable local laws, regulations (e.g., data privacy, IP, consumer protection), and contractual obligations?: Ad-hoc review process**

- **Does your organization have an incident response plan specifically addressing potential breaches, misuse, or failures related to GenAI systems, including notification procedures and mitigation strategies?: General IT incident plan adapted for AI**

- **Has a Quality Management System (QMS) or equivalent set of processes been established or adapted to oversee the lifecycle of GenAI models, including development, testing, validation, and monitoring for performance and compliance?: No specific QMS for GenAI**

- **Are employees who develop, deploy, or use GenAI systems provided with regular training on relevant policies, ethical guidelines, legal obligations, and potential risks associated with GenAI?: Ad-hoc or initial training only**

Data

- **Do GenAI systems processing personal data comply with GDPR principles (e.g., lawfulness, fairness, transparency, data minimisation)?: Partially compliant**

- **Are Data Processing Agreements (DPAs) in place with third-party GenAI providers/users involving personal data?: For some parties**

- **If using synthetic data for GenAI, has its quality, representativeness, and re-identification risks been assessed?: Aware of risks, not assessed**

- **Are mechanisms in place for data subject rights (e.g., access, rectification, erasure) for personal data used/generated by GenAI systems?: No / Not applicable**

Security

- **Are defences implemented against common adversarial attacks on GenAI models (e.g., data poisoning, model evasion, prompt injection)?: No specific defences**

- **Is data for training, fine-tuning, and inference of GenAI models secured against unauthorised access, leakage, or corruption?: Yes, strong security**

- **Are your GenAI models (weights, architecture) protected against theft or unauthorised modification?: Moderate protection**

- **Are robust access controls and authentication for users and systems interacting with GenAI applications in place?: Standard controls**

- **Is there an incident response plan specifically addressing security breaches or failures related to GenAI systems?: No specific AI incident plan**

Ethics

- **Have ethical guidelines for GenAI development and deployment been established or adopted?: Considering guidelines**

- **Are GenAI systems monitored for fairness using defined metrics, with identified disparities addressed?: Periodic monitoring**

- **Is there a clear process for human oversight and intervention in GenAI decisions/content, especially in sensitive contexts?: Limited human oversight**

- **Do you engage diverse stakeholders (including affected communities) on ethical implications of GenAI systems?: Occasional engagement**

- **Are clear accountability mechanisms in place for outcomes and decisions by or assisted by GenAI systems?: Limited accountability**

Capability & Readiness

- **Is it clearly disclosed to users when they interact with GenAI systems or consume AI-generated content (e.g., deepfakes, text)?: No disclosure**

- **Can GenAI systems provide context-appropriate explanations or justifications for their outputs/decisions?: Yes, satisfactory degree**

- **Do users have appropriate control over GenAI system operation and outputs (e.g., stop, correct, override)?: Some control**

- **Are GenAI systems tested for robustness and reliability under various conditions (including edge cases, unexpected inputs)?: Moderately tested**

- **Is comprehensive technical documentation maintained for your GenAI systems, as required by the EU AI Act for high-risk systems?: No formal technical documentation**

Assessment Summary

Overall AI Security Maturity

Defined

AI security processes are formally defined and documented. Consistent implementation across projects is the next key area of focus.

52% Complete

Key Domain Insights

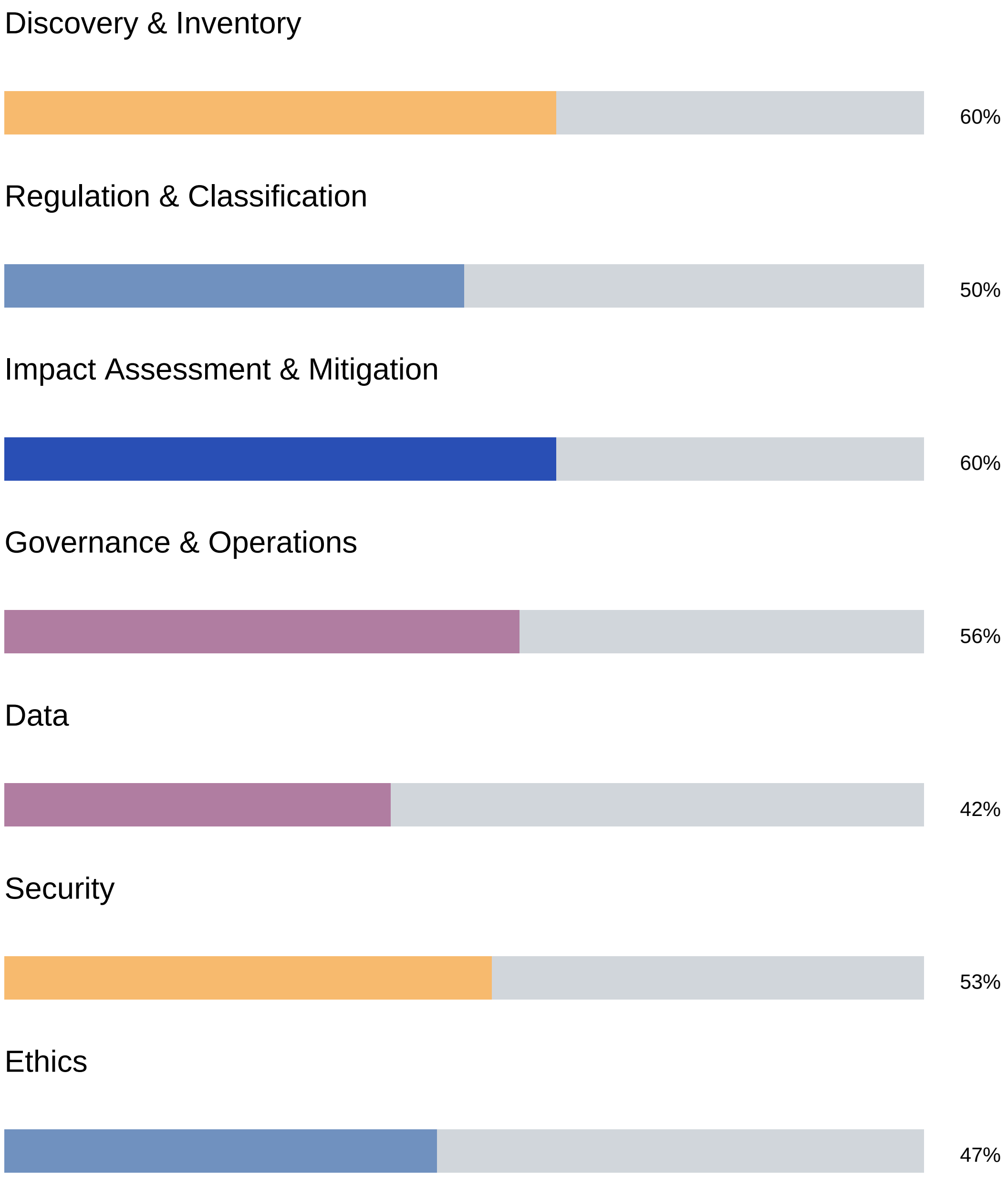
• Discovery & Inventory: 60% - Developing

• Regulation & Classification: 50% - Developing

• Impact Assessment & Mitigation: 60% - Developing

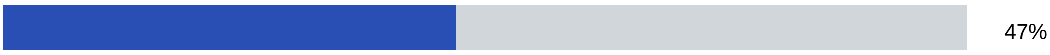
• Governance & Operations: 56% - Developing

- Data: 42% - Developing
- Security: 53% - Developing
- Ethics: 47% - Developing
- Capability & Readiness: 47% - Developing



Key Domain Insights

Capability & Readiness



Key Recommendations

Strengthen Data Governance & Privacy Frameworks

Your responses indicate potential gaps in data handling and privacy policies for AI systems. DevSecAI's Data Governance service helps establish robust frameworks, ensuring compliance with regulations like GDPR, CCPA, and managing data lifecycle for AI.

Enhance AI Model Security & Integrity

Concerns regarding AI model security, including vulnerability to adversarial attacks or lack of integrity checks, were noted. DevSecAI offers AI Model Security assessments and hardening services to protect your valuable AI assets.

Next Steps

- Schedule a comprehensive review of your AI systems
- Develop a compliance roadmap based on this assessment
- Establish documentation procedures for high-priority areas
- Review and update risk management processes
- Consider expert consultation for complex compliance requirements

Book a Consultation

For personalised guidance on implementing these recommendations and ensuring full compliance with the EU AI Act, book a consultation with our experts at:

https://www.devsecai.io/contact-us