AI Governance Desk

Data Lineage
Compliance Audit Checklist

EU AI Act–Aligned Checklist for High-Risk AI Systems

For Chief Compliance Officers, Legal & AI Governance Teams

This document supports governance readiness and does not constitute legal advice.

How to Use This Data Lineage Compliance Audit Checklist

A practical self-assessment tool for EU AI Act readiness

Purpose of This Checklist

This checklist is designed to help organizations assess whether their data lineage documentation meets the requirements of the EU AI Act for high-risk AI systems. It supports governance teams in identifying gaps, prioritizing remediation actions, and preparing for regulatory audits.

Who Should Use This Checklist

- Chief Compliance Officers (CCOs)
- Legal & Regulatory Teams
- AI Governance & Risk Committees
- Data Protection Officers (DPOs)
- Internal Audit & Risk Teams

Understanding Your Responses

- Yes – Fully implemented and documented
- Partial – Implemented but incomplete or inconsistent
- No – Not implemented or undocumented

"Partial" and "No" responses indicate potential compliance risk and should be prioritized for remediation.

How to Complete the Checklist

1. Review each question carefully
2. Select Yes, Partial, or No based on current practices
3. Add notes or evidence where applicable
4. Focus on high-risk AI systems first

Recommended Next Steps

- Prioritize remediation for high-risk systems
- Assign ownership for each identified gap
- Update documentation and controls
- Reassess periodically or after system changes

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Data Source & Provenance

EU AI Act — Article 10 (Data Governance & Training Data)

This section assesses whether the origins, legality, and relevance of data used in high-risk AI systems are clearly documented and traceable.

Are all data sources used in the AI system clearly identified and documented?
□ Yes □ Partial □ No
Notes / Evidence:

Is the original source of each dataset (internal, third-party, public, scraped) recorded?
□ Yes □ Partial □ No
Notes / Evidence:

Is the lawful basis for using each dataset documented (e.g., consent, legitimate interest, contractual use)?
□ Yes □ Partial □ No
Notes / Evidence:

Have data licensing, usage rights, and restrictions been reviewed and recorded?
□ Yes □ Partial □ No
Notes / Evidence:

Are personal data sources clearly distinguished from non-personal data sources?
□ Yes □ Partial □ No
Notes / Evidence:

Is the relevance of each dataset to the AI system's intended purpose documented?
□ Yes □ Partial □ No
Notes / Evidence:

Are known data gaps, limitations, or provenance risks identified and documented?
□ Yes □ Partial □ No
Notes / Evidence:

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Data Collection, Representativeness & Bias

EU AI Act — Article 10 (Bias, Representativeness, Errors)

This section assesses whether data used to train and validate the AI system is representative, bias-aware, and appropriate for the system's intended purpose.

Are data collection methods for the AI system clearly defined and documented?
□ Yes □ Partial □ No
Notes / Evidence:

Has the dataset been assessed for representativeness relative to the AI system's intended purpose and target population?
□ Yes □ Partial □ No
Notes / Evidence:

Have potential sources of bias in the data (e.g., demographic, geographic, temporal, or proxy bias) been identified and documented?
□ Yes □ Partial □ No
Notes / Evidence:

Are documented measures in place to mitigate identified biases in the dataset?
□ Yes □ Partial □ No
Notes / Evidence:

Has the dataset been reviewed for data quality issues such as errors, missing values, or outdated information?
□ Yes □ Partial □ No
Notes / Evidence:

Are residual bias risks or data limitations clearly documented and communicated to relevant governance or oversight teams?
□ Yes □ Partial □ No
Notes / Evidence:

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Data Preparation, Transformation & Labeling

EU AI Act – Article 10 (Data Processing & Quality Controls)

This section assesses whether data preparation, transformation, and labeling processes are clearly documented, reproducible, and designed to maintain data quality, integrity, and compliance throughout the AI lifecycle.

Are all data preparation steps (cleaning, filtering, normalization) clearly defined and documented?

□ Yes □ Partial □ No

Notes / Evidence:

Is version control applied to datasets after preparation or transformation?

□ Yes □ Partial □ No

Notes / Evidence:

Are labeling or annotation processes clearly documented, including guidelines provided to annotat

□ Yes □ Partial □ No

Notes / Evidence:

Have labeling quality controls (e.g., double labeling, inter-annotator agreement, review processes) been implemented and recorded?

□ Yes □ Partial □ No

Notes / Evidence:

Are automated preprocessing or transformation tools logged and auditable?

□ Yes □ Partial □ No

Notes / Evidence:

Are data preparation and labeling decisions reviewed for potential bias introduction or amplificatio

□ Yes □ Partial □ No

Notes / Evidence:

Are data transformation processes (e.g., feature engineering, aggregation, encoding) traceable and reproducible?

□ Yes □ Partial □ No

Notes / Evidence:

Are known limitations, assumptions, or risks introduced during preparation or labeling documented and communicated to governance teams?

□ Yes □ Partial □ No

Notes / Evidence:

Document evidence may include data pipeline documentation, version control logs, annotation guidelines, or quality assurance reports.

Data Lineage, Traceability & Documentation

EU AI Act – Articles 10 & 12 (Data Governance, Record-Keeping & Technical Documentation)

This section assesses whether end-to-end data lineage, traceability, and supporting documentation enable transparency, auditability, and regulatory accountability for high-risk AI systems.

Are complete data lineage records maintained from original data sources through model deployment?
□ Yes □ Partial □ No
Notes / Evidence:

Can datasets used for training, validation, and testing be traced back to their original sources and versions?
□ Yes □ Partial □ No
Notes / Evidence:

Are data transformations, preprocessing steps, and feature engineering activities traceable and documented?
□ Yes □ Partial □ No
Notes / Evidence:

Is lineage information maintained across internal teams, vendors, and third-party data providers?
□ Yes □ Partial □ No
Notes / Evidence:

Are changes to datasets, pipelines, or lineage records versioned and logged over time?
□ Yes □ Partial □ No
Notes / Evidence:

Can lineage documentation support incident investigation, audit requests, or regulatory inquiries?
□ Yes □ Partial □ No
Notes / Evidence:

Are lineage responsibilities clearly assigned to accountable owners or teams?
□ Yes □ Partial □ No
Notes / Evidence:

Are known lineage gaps, limitations, or traceability risks identified and documented?
□ Yes □ Partial □ No
Notes / Evidence:

Document evidence may include data lineage diagrams, pipeline logs, metadata catalogs, version histories, or audit trail records.

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Model Training, Validation & Evaluation

EU AI Act — Article 10 (Training, Validation & Testing Data) & Article 15 (Accuracy, Robustness, Cybersecurity)

This section assesses whether model training, validation, and evaluation processes are documented, controlled, and designed to ensure accuracy, robustness, fairness, and compliance with EU AI Act requirements for high-risk AI systems.

Are training, validation, and testing datasets clearly defined and documented?

□ Yes □ Partial □ No

Notes / Evidence:

Is the separation between training, validation, and test datasets clearly enforced and documented?

□ Yes □ Partial □ No

Notes / Evidence:

Are model training procedures, parameters, and configurations documented and reproducible?

□ Yes □ Partial □ No

Notes / Evidence:

Have appropriate evaluation metrics been defined based on the AI system's intended purpose and risk profile?

□ Yes □ Partial □ No

Notes / Evidence:

Has model performance been evaluated for accuracy, robustness, and potential bias across relevant groups or scenarios?

□ Yes □ Partial □ No

Notes / Evidence:

Are validation and testing results reviewed, approved, and recorded before deployment?

□ Yes □ Partial □ No

Notes / Evidence:

Are known model limitations, assumptions, or performance trade-offs documented and communicated to relevant stakeholders?

□ Yes □ Partial □ No

Notes / Evidence:

Is there a process for re-evaluating model performance after updates, retraining, or significant data changes?

□ Yes □ Partial □ No

Notes / Evidence:

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Human Oversight & Accountability (EU AI Act Article 14)

EU AI Act – Article 14 (Human Oversight)

This section assesses whether appropriate human oversight mechanisms are in place to ensure that high-risk AI systems can be effectively monitored, intervened in, and overridden to prevent or mitigate risks throughout their lifecycle.

Are clear human oversight roles and responsibilities defined for the AI system?

□ Yes □ Partial □ No

Notes / Evidence:

Are designated personnel trained and competent to understand the AI system's outputs, limitations, and risks?

□ Yes □ Partial □ No

Notes / Evidence:

Are procedures in place that allow humans to intervene, override, or stop the AI system when necessary?

□ Yes □ Partial □ No

Notes / Evidence:

Are escalation and incident response processes defined when the AI system behaves unexpectedly or produces harmful outcomes?

□ Yes □ Partial □ No

Notes / Evidence:

Are decisions influenced or made by the AI system subject to appropriate human review, especially in high-risk or sensitive contexts?

□ Yes □ Partial □ No

Notes / Evidence:

Are limitations, assumptions, and known risks of the AI system clearly communicated to human operators and decision-makers?

□ Yes □ Partial □ No

Notes / Evidence:

Is accountability for AI-related decisions clearly assigned to specific individuals, roles, or governance bodies?

□ Yes □ Partial □ No

Notes / Evidence:

Are human oversight mechanisms periodically reviewed and updated based on system changes, incidents, or regulatory developments?

□ Yes □ Partial □ No

Notes / Evidence:

Document evidence may include oversight procedures, training records, escalation protocols, decision review logs, incident reports, or governance accountability frameworks

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Risk Management & Incident Handling

Risk Management & Incident Handling

EU AI Act – Articles 9, 61 & 62

This section assesses whether a structured risk management and incident handling framework is in place to identify, assess, mitigate, monitor, and report risks associated with high-risk AI systems throughout their lifecycle.

Is a documented AI risk management process in place covering the full AI lifecycle?

□ Yes □ Partial □ No

Notes / Evidence:

Are foreseeable risks to health, safety, fundamental rights, and compliance systematically identified and assessed?

□ Yes □ Partial □ No

Notes / Evidence:

Are risk mitigation measures defined, implemented, and proportionate to the identified risks

□ Yes □ Partial □ No

Notes / Evidence:

Are residual risks evaluated and formally accepted by appropriate governance or manageme bodies?

□ Yes □ Partial □ No

Notes / Evidence:

Is there a documented incident detection and response process for AI system failures or harmful outcomes?

□ Yes □ Partial □ No

Notes / Evidence:

Are serious incidents and malfunctions recorded, investigated, and escalated in a timely manner?

□ Yes □ Partial □ No

Notes / Evidence:

Are procedures in place to notify competent authorities and relevant stakeholders when required under the EU AI Act?

□ Yes □ Partial □ No

Notes / Evidence:

Are risk management processes periodically reviewed and updated based on incidents, system changes, or regulatory updates?

□ Yes □ Partial □ No

Notes / Evidence:

Document evidence may include risk registers, impact assessments, mitigation plans, incident reports, escalation logs, or regulatory notifications.

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Post-Market Monitoring & Continuous Compliance

Post-Market Monitoring & Continuous Compliance
EU AI Act – Articles 61 & 62

This section assesses whether post-market monitoring processes are in place to continuously evaluate the performance, safety, and compliance of high-risk AI systems after deployment, including mechanisms for detecting issues, responding to incidents, and implementing corrective actions.

Is a documented post-market monitoring plan in place for the AI system?
□ Yes □ Partial □ No
Notes / Evidence:

Are AI system performance, accuracy, and behavior monitored on an ongoing basis after deployment?
□ Yes □ Partial □ No
Notes / Evidence:

Are mechanisms in place to detect performance degradation, unexpected behavior, or emerging risks in real-world use?
□ Yes □ Partial □ No
Notes / Evidence:

Are user feedback, complaints, or adverse events systematically collected and reviewed?
□ Yes □ Partial □ No
Notes / Evidence:

Are post-deployment findings documented and assessed for compliance impact?
□ Yes □ Partial □ No
Notes / Evidence:

Are corrective or preventive actions implemented when issues or non-compliance are identified?
□ Yes □ Partial □ No
Notes / Evidence:

Are significant system changes, updates, or retraining activities evaluated for renewed compliance risk?
□ Yes □ Partial □ No
Notes / Evidence:

Are serious incidents reported to competent authorities within required EU AI Act timelines?
□ Yes □ Partial □ No
Notes / Evidence:

Document evidence may include post-market monitoring plans, performance dashboards, incident logs, user feedback records, corrective action reports, or regulatory submissions.

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Overall AI Compliance Summary

This summary provides an overall view of the organization's data lineage and AI governance readiness based on responses in this checklist. "Partial" and "No" responses indicate areas requiring remediation before EU AI Act compliance can be demonstrated.

- **Number of Yes responses: _____**
- **Number of Partial responses: _____**
- **Number of No responses: _____**

Risk Rating (pick one):

□ **Low Risk**
□ **Medium Risk**
□ **High Risk**

Risk rating is determined based on the volume and severity of "Partial" and "No" responses.

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Required Remediation & Action Plan

- **Identified Gap:** _____
- **Affected Section(s):** _____
- **Required Action:** _____
- **Responsible Owner:** _____
- **Target Completion Date:** _____

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Declaration of Accountability

We confirm that this checklist has been completed accurately and in good faith to assess compliance with EU AI Act requirements for high-risk AI systems. Identified gaps will be addressed in accordance with organizational governance processes.

- Name: _____
- Role / Title: _____
- Organization: _____
- Date: _____
- Signature: _____

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

Appendix – Supporting Documentation Index

This appendix provides a non-exhaustive index of documents that may be referenced as evidence to support responses recorded in this checklist. Availability of documents may vary depending on the AI system's risk profile, deployment context, and organizational maturity.

- Data source inventories and provenance records
- Dataset documentation and data sheets
- Data preprocessing and transformation logs
- Annotation guidelines and labeling instructions
- Version control logs (datasets and models)
- Model training and evaluation reports
- Bias, fairness, and robustness assessment reports
- Risk assessments and risk registers
- Human oversight procedures and escalation protocols
- Incident logs and post-incident review reports
- Post-market monitoring plans and performance dashboards
- Regulatory notifications and authority correspondence

AI Governance Desk | This document supports governance readiness and does not constitute legal advice.

AI Governance Desk

AI Governance Desk provides practical governance tools, compliance frameworks, and readiness assessments to support organizations deploying or managing high-risk AI systems under the EU AI Act and other emerging global AI regulations.

Why This Checklist Matters

This checklist is designed to help organizations:

- Assess readiness for EU AI Act compliance
- Identify gaps in data governance, risk management, and human oversight
- Support internal audits and regulatory inspections
- Improve accountability, transparency, and operational control of AI systems
- Reduce legal, ethical, and reputational risks associated with high-risk AI deployments

Who Should Use This Document

- AI product owners and system developers
- Compliance and risk management teams
- Legal and regulatory affairs professionals
- Data protection officers and governance leads
- Internal audit and assurance teams
- Executive leadership overseeing AI strategy