



Generative AI Policy

DOC-IT-IS-27
Revision 1.0
NOVEMBER 1, 2025

REVISION HISTORY

This revision history lists all changes to this document, the date of each revision, and the reason for each revision.

<u>Rev.</u>	<u>Rev. Date</u>	<u>Revision Description</u>
1.0	11/01/2025	Initial Version

27 Generative AI Policy	4
27.1 Executive Summary	4
27.2 Purpose	4
27.3 Scope	4
27.4 Definitions.....	4
27.5 AI Policy Requirements	5

27 Generative AI Policy

27.1 Executive Summary

Generative AI tools including but not limited to ChatGPT, Bard, Jasper AI, Bing AI, CoPilot, and any Large Language Model (LLM) or agentic AI system offer significant opportunities for innovation and efficiency. However, their use introduces risks related to cybersecurity, compliance, privacy, intellectual property, and ethical outcomes. This policy establishes RXO's standards for responsible, secure, and compliant use of generative AI. The standard applies to all systems, applications, and network devices managed by RXO, including on-premises infrastructure and cloud environments. It supports the broader goals of the Information Security Management System (ISMS) and aligns with frameworks such as ISO/IEC 27001, NIST SP 800-53, and CIS Controls.

27.2 Purpose

The purpose of this Policy is to establish the Generative AI requirements governing the practices and processes incorporated within RXO's Cybersecurity Program and Secure Computing Environment implemented for storing, processing, and transmitting sensitive company and customer data. This Policy is intended to outline Generative AI Practices and the Minimum Standard for protecting RXO and customer data, and the systems which store, process, and transmit data within RXO's system boundaries.

27.3 Scope

This standard applies to RXO, Inc., including all subsidiaries, divisions, and operating entities. All directors, officers, employees, contractors, and third parties acting on behalf of RXO must comply with this policy.

27.4 Definitions

Term	Definition
Generative AI	Any system capable of producing text, code, images, or other content using machine learning models.
Public AI	AI tools or platforms accessible to the general public.
Private AI	AI tools deployed within RXO's secure infrastructure.
Third-Party AI	AI tools or services provided by external vendors.
Enterprise AI	AI systems developed, owned, or operated by RXO.

27.5 AI Policy Requirements

27.5.1 General Prohibitions

No RXO proprietary, personal, regulated, or confidential information including HIPAA, PII, or other sensitive data may be entered into any public or third-party generative AI tool unless the tool is explicitly certified for such use and written approval is obtained from Information Security. Customer contracts may contain specific prohibitions on using their data in AI platforms or to train AI models, and those contract provisions must be always followed.

27.5.2 Approved AI Platforms

Only the Microsoft Copilot 'Work' version is authorized for use by RXO personnel. Use of any other generative AI platform including public, third-party, or non-Microsoft solutions requires prior written approval from the Information Security team.

27.5.3 Use of AI Output

AI-generated output may not be incorporated into RXO products, deliverables, innovations, or intellectual property including source code, documentation, or operating procedures without express written authorization from Information Security. AI-generated output must be validated by a qualified human reviewer before use in RXO business decisions. Remediate findings within 30 days or submit a risk acceptance form to the BISO.

27.5.4 Building & Utilizing Generative AI Tools

Any project involving the building, integration, or use of generative AI tools requires consultation with the Architecture and Infosec teams and written approval from Information Security.

27.5.5 Data Security & Privacy

Data Protection Impact Assessments (DPIA) are required for any new AI deployment. All data processed by AI must be encrypted at rest and in transit. Retention and deletion policies must be defined and enforced.

27.5.6 Model Integrity & Risk Management

All AI models must be protected against adversarial attacks and input manipulation. Regular audits for bias, accuracy, fairness, and drift are required.

27.5.7 Governance & Compliance

This policy aligns with NIST AI Risk Management Framework (RMF), ISO/IEC 42001, and other relevant standards. All AI projects must maintain documentation, versioning, and audit trails.

27.5.8 Safeguards

To strengthen RXO's AI risk posture, the following safeguards are mandated to ensure critical threats are checked:

Threat	Statement	Key Safeguard
Data Poisoning	R XO datasets cannot be compromised	Dataset provenance checks, DLP scans, controlled ingestion pipelines
Prompt Injection	Malicious prompts cannot alter AI behavior	Prompt sanitization, middleware isolation, WAF rules
Model Theft	R XO models cannot be exfiltrated or cloned	IAM-based authentication, rate limiting, anomaly detection
Jailbreaks	Safety filters cannot be bypassed	Response classification, safety filters, restricted system prompts
Hallucinations	AI cannot produce ungrounded outputs	Grounding APIs, fact-checking, human validation

27.5.9 Policy Integration

- Data Security & Privacy: Extend to include dataset validation and provenance tracking.
- Model Integrity & Risk Management: Add adversarial resilience and anomaly detection requirements.
- Use of AI Output: Require grounding responses to verified RXO knowledge bases.
- Governance & Compliance: Mandate documentation of safeguards and periodic audits.
- Ethical Use & Transparency: Reinforce fairness and non-discrimination with jailbreak prevention measures.

27.5.10 Technical Implementation

- Apply GCP/Vertex AI controls for prompt sanitization, safety filters, and grounding.
- Maintain logs for all safeguard enforcement actions.
- Include these safeguards in DPIAs and annual compliance reviews. NSA Network Infrastructure Security Guide

Certificate of Review

Generative AI Policy

December 18th 2025 10:42:25 am

Name	Job title	Stakeholder type	Action	Timestamp
Mark Gutman		Policy Owner	Submitted	Friday, November 21, 2025 11:22 AM EST
Faisal Khan		Policy Approver	Approved	Friday, November 21, 2025 12:12 PM EST
Mark Gutman		Policy Admin	Published	Thursday, December 18, 2025 10:42 AM EST