Compliance evidence — agentic AI security assessment

> Sample. The target is fictional and the scan is the product's demo. This document is rendered by the same code that renders a customer's download, with the same redaction.

System under test: Northwind Health Plus Chatbot
Assessment: demo-customer-support
Date: 2026-03-22T05:00:13Z
Frameworks: EU AI Act (Regulation (EU) 2024/1689), GDPR (Regulation (EU) 2016/679), OWASP LLM Top 10, MITRE ATLAS

> This document is the technical evidence produced by adversarial testing. It is an input to a conformity assessment or a breach assessment, not the assessment itself, and it carries no provider sign-off.

> **Scan profile:** deep.

Position

Techniques executed	352
Findings	7

Article-by-article analysis

Article	Obligation	Findings	GDPR	Breach clock
Art10	Data and Data Governance	5	Art17, Art32, Art33, Art44, Art5, Art6	yes — GDPR Art.33 breach-notification assessment starts
Art13	Transparency and Provision of Information to Deployers	5	Art17, Art32, Art33, Art44, Art5, Art6	yes — GDPR Art.33 breach-notification assessment starts
Art14	Human Oversight	1	Art25, Art32, Art33	yes — GDPR Art.33 breach-notification assessment starts
Art15	Accuracy, Robustness, and Cybersecurity	6	Art17, Art25, Art32, Art33, Art44, Art5, Art6	yes — GDPR Art.33 breach-notification assessment starts
Art9	Risk Management System	1	Art25, Art32, Art33	yes — GDPR Art.33 breach-notification assessment starts

GDPR articles implicated by findings that map to no EU AI Act obligation: Art17, Art44, Art5, Art6

At least one finding starts a GDPR Article 33 assessment. The 72-hour clock runs from awareness, and this document is a record of awareness.

Obligations not assessed

These articles were not put under test by this assessment. They are neither passed nor failed: no evidence was produced about them.

Article	Obligation	Why it is not assessed
Art5	Prohibited AI Practices	no technique executed under the ART5 category, the only route to it — and Article 5 has applied since 2 February 2025, so this gap is live, not deferred to 2027

Article	Obligation	Why it is not assessed
Art6	Classification Rules for High-Risk AI Systems	no technique executed under a category that implicates it
Art12	Record-Keeping	no technique on a route to it executed — it is reached only by OWASP-Agentic T8 and OPENCLAW.log.poisoning
Art50	Transparency Obligations for Certain AI Systems	no technique on a route to it executed — it is reached only by OWASP-Agentic T15

GDPR exposure

Article	Obligation	Findings	Breach-notification candidate
Art17	Right to Erasure (Right to be Forgotten)	6	**yes**
Art25	Data Protection by Design and by Default	1	**yes**
Art32	Security of Processing	6	**yes**
Art33	Notification of a Personal Data Breach to the Supervisory Authority	6	**yes**
Art44	General Principle for International Data Transfers (Chapter V)	6	**yes**
Art5	Principles Relating to Processing of Personal Data	6	**yes**
Art6	Lawfulness of Processing	6	**yes**

At least one finding exposed personal data in a way that may engage the Article 33 duty to notify the supervisory authority within 72 hours. Whether the duty is actually engaged is a determination for the controller: this assessment establishes that the exposure is reachable, not that a breach of personal data has occurred.

Regulatory Crosswalk

Evidence toward, not a test of. These frameworks are audited at the organisation and lifecycle level; this assessment is one technical input to the named control, not a substitute for the audit.

This assessment delivered 352 adversarial techniques to the target and recorded 7 finding(s). On that basis it is evidence toward the following controls:

Framework	Control	What it requires	Basis
DORA	Art. 24-27	Threat-led penetration testing and advanced testing of ICT tools and systems	Adversarial testing of an AI agent is a component of the TLPT scope for that ICT system.
DORA	Art. 8-9	ICT risk identification and protection	Findings identify exploitable ICT risks in a deployed AI system and their protective gaps.
NIS2	Art. 21(2)(e)	Security in acquisition, development and maintenance, including vulnerability handling	The assessment is the vulnerability-testing input to this risk-management measure for AI systems.
NIS2	Art. 21(2)(a)	Policies on risk analysis and information system security	Coverage and findings substantiate that AI-specific risks were analysed, not assumed.
ISO 42001	A.6.2.4	AI system verification and validation	An adversarial assessment is verification-and-validation evidence for the AI system control.
ISO 27001	A.8.8	Management of technical vulnerabilities	Findings are the identified technical vulnerabilities this control requires be managed.
ISO 27001	A.8.29	Security testing in development and acceptance	The scan is security-testing evidence for the AI component before or after acceptance.

Framework	Control	What it requires	Basis
CRA	Annex I, Part I(2)	No known exploitable vulnerabilities at time of placing on the market	An assessment that finds none, or that lists what was fixed, is evidence toward this essential requirement.
CRA	Annex I, Part II	Vulnerability handling requirements	Findings feed the documented handling process the manufacturer must operate.
CSA MAESTRO	Layer 3 — Agent Frameworks & Tools	Threat modelling of the tool surface an agent can reach	An adversarial assessment exercises the tool surface directly and is evidence for this layer's model.
CSA MAESTRO	Layer 5 — Evaluation & Observability	Continuous evaluation of agent behaviour under adversarial input	Repeated assessments across releases are the evaluation evidence this layer asks for.
NIST AI RMF	MEASURE 2.7	AI system security and resilience are evaluated and documented	The assessment is the evaluation, and its report is the documentation this subcategory requires.
NIST AI RMF	MANAGE 4.1	Post-deployment monitoring plans are implemented, including for AI system risks	Re-running the assessment against a deployed agent after each release is one implemented mechanism.
FINMA	Cyber / ICT risk	Testing of the security of critical ICT functions	For a supervised institution, the assessment evidences testing of a critical AI-driven function.