European Journal of Risk Regulation (2026), 1–10
doi:10.1017/err.2026.10159

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Persistent Agent Memory under the EU AI Act and NIS2: Misalignment, Emergent Behaviour and Regulatory Fit

Albert Adusei Brobbey

Crestmont University, Irvine, USA
Email: aduseialbert@rocketmail.com

Abstract

Agentic AI systems commonly combine a general-purpose AI model with prompts, tools, orchestration and an external memory layer. The model need not be fine-tuned: records retrieved from persistent memory are normally inserted into the model's inference-time context and can thereby influence later outputs and actions without changing model weights. This article asks how far the EU Artificial Intelligence Act and Directive (EU) 2022/2555 (NIS2) already govern the risks that follow. It reframes persistent memory as a system feature that can mediate or amplify misalignment and emergent behaviour, rather than as training data or a new category in the AI Act's use-case-based taxonomy. Article 15 provides a clear legal anchor for robustness and cybersecurity, while Articles 9–14, 25, 72 and 73 extend coverage across risk management, data governance, logging, oversight, value-chain responsibility and post-market monitoring. For organisations within NIS2's scope, Articles 20, 21 and 23 add entity-level governance, all-hazards cybersecurity, supply-chain controls and incident reporting. The remaining problem is therefore not a wholesale absence of law. It is whether these overlapping duties are sufficiently operationalised for memory-mediated failures such as poisoned retrieval, scope leakage, stale reliance, revocation failure and loss of provenance. The article proposes proportionate implementation through existing guidance, standards, documentation and monitoring mechanisms, without creating a new statutory risk tier or a memory-specific regulatory regime.

Keywords: agentic AI; cybersecurity; emergent behaviour; EU AI Act; NIS2; persistent memory

I. Introduction

The EU Artificial Intelligence Act is a risk-based instrument whose principal obligations turn on prohibited practices, high-risk use cases, transparency situations and general-purpose AI (GPAI) models.[1] Agentic systems complicate this architecture because they are usually composite deployments rather than a single model. A typical agent combines a standard front-end large language model, system and user prompts, retrieval, persistent storage, tools, permissions and workflow logic. The underlying model may be a GPAI model, but the assembled agent is ordinarily assessed as an AI system by reference to its intended purpose and context of use. Connecting a GPAI model to employment, credit,

[1] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) [2024] OJ L, 12 July 2024 (AI Act).

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/err.2026.10159

health or public-service records can therefore create a high-risk system even though the base model remains general purpose.[2]

This article examines one component of that assembled system: persistent memory. Persistent memory can give an agent continuity across sessions and tasks, but it can also preserve false premises, unauthorised instructions, outdated policies and improperly scoped personal or organisational information. When retrieved, those records can alter the context in which the model selects a response or tool action. The regulatory concern is consequently not memory in isolation. It is the way memory interacts with model behaviour, tool access, institutional objectives and the deployment environment.

The argument is deliberately narrower than the original proposition that persistent memory should be recognised as a new regulatory risk category. The AI Act's categories are principally context- and use-specific. Persistent memory does not displace that architecture and should not become a fifth tier. It is better understood as a risk-relevant system feature: one that may create, transmit or intensify misalignment and emergent behaviour within an existing legal classification. The central question is therefore one of regulatory fit. To what extent do the AI Act and NIS2 already cover memory-mediated risk, and where do their general requirements need more precise implementation?

The answer is that the coverage is substantial. The AI Act already requires lifecycle risk management, data governance, logging, transparency, human oversight, accuracy, robustness and cybersecurity for high-risk systems. Article 15 expressly covers resilience to errors, faults, inconsistencies, interactions with people or other systems, unauthorised attempts to alter performance and appropriate technical measures against AI-specific vulnerabilities. NIS2 separately regulates the network and information systems of essential and important entities through management accountability, all-hazards risk management, incident handling, business continuity, supply-chain security, access control and incident reporting.[3]

The contribution is thus diagnostic and implementation-oriented. First, it separates inference-time memory from training and fine-tuning. Second, it places memory within the literature on agentic misalignment and system-level emergent behaviour. Third, it maps the relevant risks onto the AI Act and NIS2 without exaggerating gaps. Finally, it identifies residual issues – particularly memory provenance, retrieval scope, revocation, freshness and decision-context evidence – that can be addressed through existing guidance, harmonised standards, technical documentation, contractual controls and post-market monitoring.

II. From model output to agent behaviour: training, memory and misalignment

Persistent agent memory means stored or retrievable information that an AI system can use across interactions to influence later outputs, recommendations, decisions or actions. It may be implemented through a vector database, structured profile, episodic record, summary, embedding index, cache, tool-state store or cloud-edge synchronisation layer. This functional definition follows the architecture of modern agents, in which memory modules retrieve selected prior information to support planning and action.[4]

[2] AI Act, Arts. 3(1), 3(63), 6, 25 and Annex III.

[3] Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union [2022] OJ L333/80 (NIS2), Arts. 20, 21 and 23.

[4] J. S. Park and others, "Generative Agents: Interactive Simulacra of Human Behavior" (UIST 2023) 1; N. Shinn and others, "Reflexion: Language Agents with Verbal Reinforcement Learning" (NeurIPS 2023) 8634; W. Xu and others, "A-Mem: Agentic Memory for LLM Agents" (NeurIPS 2025).

https://doi.org/10.1017/err.2026.10159

This memory is not ordinarily training data. Pre-training and fine-tuning change model parameters through an optimisation process. By contrast, a standard agent can retrieve a record after deployment and place it into the prompt or context supplied to an unchanged model at inference. The generated output is then conditional on model weights and on the immediate inference context, which may contain user instructions, system instructions, retrieved memory, tool results and conversation history. Persistent storage and the context window should also be distinguished: the store survives across invocations, while selected records are copied into the finite context used for a particular inference. Only where retained interactions are later used for fine-tuning or continual learning do they also become training data.

The relevant safety concept is misalignment: divergence between the behaviour produced by the agent and the objectives, constraints or values that the user, deploying organisation or law requires. Increasing autonomy, goal-directedness, long-horizon planning and direct access to tools can widen the consequences of such divergence.[5] Controlled research has shown that models placed in simulated agent roles can select harmful actions when assigned goals conflict with organisational objectives or continued operation, although the researchers expressly did not claim that such conduct had been observed in real deployments.[6] This evidence supports testing and caution; it does not justify treating every agent as inherently misaligned.

Emergent behaviour is used here in a similarly restrained sense. It refers to behaviour that arises from interactions among model, memory, prompts, tools, permissions, other agents and environment and is not adequately predicted by evaluating the base model alone. The literature disputes whether some apparent capability jumps are genuinely discontinuous or artefacts of measurement, so "emergent" should not become a synonym for mysterious or unforeseeable.[7] System-level studies nevertheless identify failures arising from inter-agent misalignment, coordination and task verification even where component models are unchanged.[8]

Memory connects these concepts. A remembered instruction can preserve a proxy goal after the initiating interaction; a summary can remove qualifications that previously constrained action; repeated retrieval can reinforce a mistaken premise; and cross-agent synchronisation can propagate it. Memory does not uniquely cause misalignment, and a memoryless agent may still behave unexpectedly. Its regulatory significance is that it can make misalignment persistent, path-dependent, personalised and difficult to reconstruct. That is the article's narrower technical claim.

III. Memory-mediated risk

Six recurrent mechanisms illustrate how memory can mediate agent risk. Memory poisoning occurs when false, malicious or unauthorised content is written and later treated as trusted context. Cross-context retrieval occurs when information collected for one person, role, purpose or organisation is retrieved in another. Revocation failure occurs when corrected, expired, erased or restricted material remains operationally available. Stale reliance occurs when old law, policy or factual context shapes a current action.

[5] A. Chan and others, "Harms from Increasingly Agentic Algorithmic Systems" (FAccT 2023) 651.

[6] A. Lynch and others, "Agentic Misalignment: How LLMs Could Be Insider Threats" (Anthropic Research, 20 June 2025) <https://www.anthropic.com/research/agentic-misalignment>.

[7] J. Wei and others, "Emergent Abilities of Large Language Models" (2022) Transactions on Machine Learning Research; R. Schaeffer, B. Miranda and S. Koyejo, "Are Emergent Abilities of Large Language Models a Mirage?" (NeurIPS 2023) 55565.

[8] M. Cemri and others, "Why Do Multi-Agent LLM Systems Fail" (NeurIPS 2025, Datasets and Benchmarks Track).

https://doi.org/10.1017/err.2026.10159

Table 1. Memory-mediated risk mechanisms and controls.

Mechanism	Possible misalignment or emergent pathway	Illustrative control
Memory poisoning	Untrusted content persists and steers later reasoning or tool use.	Write authorisation, trust labels, validation, anomaly detection and quarantine.
Cross-context retrieval	A relevant record is retrieved outside its lawful person, role, purpose or organisation scope.	Identity, purpose and authority tags enforced at retrieval.
Revocation failure	Corrected, erased, expired or restricted material continues to influence behaviour.	Revocation registry, deletion propagation and invalidation of derivatives.
Stale reliance	An old policy or fact becomes a current operational premise.	Freshness thresholds, version binding and blocking for consequential uses.
Provenance collapse	A summary or embedding removes source, authority or uncertainty.	Lineage metadata, transformation records and source links.
Forensic opacity	The decision context cannot be reconstructed after harm.	Retrieval-event and tool-action logs with proportionate retention.

Provenance collapse occurs when summaries, embeddings or derived records cannot be traced to their sources. Forensic opacity occurs when investigators cannot reconstruct which records influenced an output or action.

These mechanisms are technically plausible and increasingly evidenced. Query-only memory injection research demonstrates that an attacker may influence an agent's memory without direct database access, while recent government work on agent security identifies memory and context poisoning, indirect prompt injection, specification gaming and misaligned objectives as risks requiring adapted cybersecurity practice.[9] The legal analysis should nevertheless distinguish established mechanisms from deployment prevalence. Experimental attack success rates are not estimates of real-world incidence.

Memory-mediated harm depends on context. Remembering a display preference in a writing tool is not equivalent to retaining clinical history for triage or employment records for candidate screening. Relevant variables include the consequentiality of the task, sensitivity and authority of the information, write access, duration, transformation, retrieval scope, tool permissions, human review, reversibility and the availability of evidence and redress. This contextual assessment fits the AI Act's architecture better than a categorical rule based on the mere presence of memory. Table 1 summarises the main memory-mediated mechanisms and illustrative controls.

IV. Coverage under the AI Act: broad mandates, specific questions

The AI Act does not classify systems by memory architecture. Article 6 and Annex III instead focus on product safety functions and specified high-risk uses. That is not a defect: a functional, context-specific approach prevents a harmless memory feature from attracting the same duties as memory used in employment, essential services or critical infrastructure. Article 25 also helps allocate responsibility where a downstream actor substantially modifies a high-risk system or changes an intended purpose. Adding a

[9] S. Dong and others, "Memory Injection Attacks on LLM Agents via Query-Only Interaction" (NeurIPS 2025); J. Riggs and others, Summary Analysis of Responses to the Request for Information Regarding Security Considerations for AI Agents, NIST AI 800-5 (2026).

https://doi.org/10.1017/err.2026.10159

memory layer to a GPAI-based assistant may be legally important because the integrator, not the GPAI model provider, controls the records, retrieval rules, tools and domain purpose that generate the risk.

Article 9 is the primary lifecycle anchor. Its continuous and iterative process must identify known and reasonably foreseeable risks, evaluate foreseeable misuse, test risk-management measures and incorporate post-market information. Memory poisoning, scope leakage and stale reliance can therefore be treated as risk scenarios without legislative amendment. Article 10 is relevant where datasets govern the development of a high-risk system and, under Article 10(2), requires attention to origin, collection, preparation, updating, assumptions, availability and bias. Not every operational memory record is a training, validation or testing dataset, however. The more secure doctrinal route is to use Article 9 together with Articles 12–15 for inference-time memory, rather than stretching the term "training data."

Articles 12 and 19 support traceability, but adequate logging is contextual. Where memory materially influences a consequential output, useful evidence may include the identifiers or categories of retrieved records, their scope, freshness and revocation status and resulting tool actions. Article 13 requires sufficient information for deployers to interpret outputs and use the system appropriately; Article 14 requires effective oversight. Those duties can support disclosure of memory behaviour and controls without requiring disclosure of model chain-of-thought or indiscriminate retention of sensitive content.

Article 15 directly answers the claim that the Act lacks a robustness or cybersecurity mandate. High-risk systems must maintain appropriate accuracy, robustness and cybersecurity throughout their lifecycle. They must be resilient to errors, faults and inconsistencies within the system or its environment, including interaction with people or other systems, and resilient against unauthorised attempts to alter use, outputs or performance. Paragraph 5's examples – data poisoning, model poisoning, adversarial examples, confidentiality attacks and model flaws – are non-exhaustive because technical solutions are required "where appropriate" for AI-specific vulnerabilities. A malicious memory write, poisoned retrieval or cross-user leakage can therefore fall within Article 15 even though it is not an attack on training data.[10]

The fit-for-purpose issue is narrower: Article 15 does not say how providers should test a memory pipeline, what evidence must link a retrieved record to an action, how revocation should propagate to summaries and replicas, or how responsibility should be divided among model provider, agent integrator, deployer and cloud service. Those are implementation questions for standards, technical documentation, contractual allocation and supervision. Articles 72 and 73 then extend the analysis into post-market monitoring and serious-incident reporting, while Article 86 may require an affected person to receive a clear and meaningful explanation of the AI system's role in an Annex III decision. Memory provenance can be necessary evidence for those processes, even though Article 86 does not create a right to internal reasoning traces. Table 2 maps the relevant AI Act provisions to their residual implementation questions.

V. NIS2 and the deployment environment

The AI Act does not operate alone. NIS2 applies to essential and important entities identified through its sectoral and size-based scope and national transposition; it does not apply to every organisation that deploys an agent. Where it does apply, it regulates the security of the entity's network and information systems and thus the environment in

[10] AI Act, Art. 15(1), (4) and (5), and recitals 76–7.

https://doi.org/10.1017/err.2026.10159

Table 2. AI Act coverage and residual implementation questions.

AI Act provision	Existing coverage	Residual fit-for-purpose question
Arts. 6, 25; Annex III	Classifies consequential uses and reallocates provider responsibility after substantial modification or changed purpose.	When does adding or reconfiguring memory change intended purpose or constitute substantial modification?
Art. 9	Continuous lifecycle risk management covers foreseeable memory-mediated failures and misuse.	Which memory scenarios, metrics and residual-risk thresholds should testing include?
Art. 10	Governs data quality and governance for relevant development datasets.	Operational memory is not automatically training, validation or testing data; other duties must carry part of the load.
Arts. 12–14	Logging, deployer information and human oversight support traceability and control.	What retrieval evidence and user-facing memory information are proportionate?
Art. 15	Clear lifecycle mandate for robustness, cybersecurity, interaction risks and resistance to manipulation and confidentiality attacks.	How should standards test poisoning, scope leakage, revocation and recovery in memory pipelines?
Arts. 72–3	Post-market monitoring and serious-incident reporting capture failures after deployment.	Which memory anomalies are useful leading indicators and when is an incident significant?

which an agent's memory store, retrieval service, identity layer and tools operate.[11] Many deployments in energy, transport, health, public administration, digital infrastructure, cloud computing, managed services and other covered sectors may therefore attract both regimes.

Article 20 places approval and oversight of cybersecurity risk-management measures at management-body level. Article 21 requires appropriate and proportionate technical, operational and organisational measures based on an all-hazards approach. Its minimum subjects include risk analysis, incident handling, business continuity, supply-chain security, secure acquisition and maintenance, effectiveness assessment, cyber hygiene, cryptography, human-resources security, access control, asset management and, where appropriate, multi-factor authentication. Article 23 adds staged notification duties for significant incidents. These provisions directly address many organisational conditions that determine whether a poisoned or unauthorised memory record can enter, persist, spread or cause service disruption.

Commission Implementing Regulation (EU) 2024/2690 supplies more detailed requirements for specified digital infrastructure, ICT service management and digital-provider entities. Its annex requires documented risk assessment and treatment, incident handling, supply-chain policies and contractual security clauses, asset inventories and classification, access control, logging and monitoring, backup, vulnerability handling and secure development.[12] A persistent memory store may accordingly be treated as an information asset, while an external model, vector database or orchestration provider may be managed as a supplier.

The regimes are complementary rather than duplicative. The AI Act attaches requirements to particular AI systems and actors and protects health, safety and

[11] NIS2, Arts. 2–3 and Annexes I–II.

[12] Commission Implementing Regulation (EU) 2024/2690 of 17 October 2024 laying down rules for the application of NIS2 [2024] OJ L, Annex, especially points 2, 3, 5, 6, 9 and 12.

https://doi.org/10.1017/err.2026.10159

Table 3. Complementary functions of the AI Act and NIS2.

Issue	AI Act contribution	NIS2 contribution for in-scope entities
Risk governance	System-level lifecycle risk management and conformity.	Entity-level, management-approved all-hazards cybersecurity risk management.
Technical resilience	Accuracy, robustness and cybersecurity of high-risk AI systems.	Security of network and information systems, access, assets, development and maintenance.
Value chain	Provider, deployer and downstream responsibility, including substantial modification.	Supplier selection, contractual clauses, monitoring and supply-chain security.
Monitoring and incidents	Post-market monitoring, corrective action and serious-incident reporting.	Detection, handling, business continuity and staged reporting of significant incidents.
Rights and explanation	Fundamental-rights safeguards, transparency, oversight and limited explanation duties.	No substitute for AI-specific rights safeguards; focus is service and information-system security.

fundamental rights. NIS2 attaches governance and cybersecurity duties to in-scope entities and protects the availability, authenticity, integrity and confidentiality of network and information systems and services. The same event may engage both: a poisoned memory may reveal non-conformity with Article 15, a weakness in Article 21 controls and, if impact thresholds are met, reporting duties under Article 73 of the AI Act and Article 23 of NIS2. Coordinated internal escalation is therefore essential, but the legal triggers and recipients remain distinct.

Table 3 compares the complementary functions of the AI Act and NIS2 for entities within NIS2's scope. NIS2 also has limits. It cannot replace the AI Act's use-case classification, fundamental-rights safeguards, conformity assessment, deployer information or explanation mechanisms. Nor does cybersecurity exhaust memory-mediated risk: a perfectly authenticated but outdated or contextually inappropriate record may produce harmful action without a security breach. Conversely, the AI Act should not be interpreted as a complete enterprise-security code. Recent scholarship appropriately treats NIS2 and related horizontal regimes as part of the assessment of reasonable cybersecurity and residual harm.[13]

VI. A combined fit-for-purpose assessment

A combined compliance assessment should begin with legal scope rather than architecture. The organisation should identify the agent's intended purpose, whether it falls within Article 6 or Annex III, which actor is the provider or deployer, whether memory integration changes intended purpose or substantially modifies the system, and whether the organisation is an essential or important entity under the applicable national NIS2 rules. Only then should it map memory components and flows.

The technical map should identify what is retained; who and what can write, retrieve, transform and delete it; where copies and derived summaries exist; which identities, roles and purposes delimit retrieval; how external content enters; which tools can act on retrieved material; and which evidence is retained. This map supports Article 9 risk

[13] F Teichmann, "Risk, Reasonableness and Residual Harm under the EU AI Act: A Conceptual Framework for Proportional Ex-Ante Controls" (2026) European Journal of Risk Regulation, https://doi.org/10.1017/err.2025.10077.

https://doi.org/10.1017/err.2026.10159

analysis, Article 13 instructions, Article 14 oversight, Article 15 testing and, where applicable, NIS2 asset, access, supplier and incident controls.

Testing should focus on behaviour, not merely the integrity of storage. Relevant tests include adversarial memory writes, retrieval across identity and purpose boundaries, conflicts between current instructions and remembered goals, revoked or stale records, loss of source qualifiers during summarisation, cloud-edge synchronisation conflicts, tool use following poisoned retrieval and recovery after quarantine. The outcome measure is whether the assembled system remains within its intended and lawful operating envelope. This is where misalignment and emergent behaviour become operational regulatory concepts rather than abstract labels.

Evidence should be proportionate. High-impact uses may justify record identifiers, retrieval scores, scope decisions, policy versions, tool calls and human interventions. Low-risk personalisation may justify far less. Logs should not become an uncontrolled secondary memory containing unnecessary personal data or secrets. Data minimisation, retention limits, access control and security must therefore apply to the evidence layer itself.

VII. Proportionate implementation within existing law

The analysis supports targeted implementation rather than a new regulatory regime. First, Commission guidance under Article 96 can explain that inference-time memory may form part of the high-risk system's operational context and should be addressed under Articles 9 and 12–15 where it materially influences outputs or actions. Guidance can also clarify the division of responsibility where a downstream integrator adds persistent memory to a GPAI-based application.

Second, harmonised standards and common specifications can convert Article 15's broad mandate into testable outcomes: authorised writes, isolation among persons and purposes, provenance of derived records, effective revocation, freshness controls, monitoring, quarantine, recovery and evidence of retrieval. These are functional requirements and therefore compatible with technology neutrality. They should be calibrated to risk and should not prescribe a particular vector database or agent framework.[14]

Thirdly, Annex IV technical documentation and Article 13 instructions can include a concise memory profile where memory is material. The profile should describe memory categories, sources, transformations, retention, write authority, retrieval boundaries, revocation, logging, known limitations and deployer controls. For NIS2 entities, the same profile can feed asset inventories, supplier requirements, incident playbooks and business continuity plans. Reusing evidence across regimes reduces compliance duplication.

Fourthly, post-market monitoring should use memory anomalies as leading indicators: repeated blocked cross-scope retrieval, access to expired records, unexplained writes, abnormal propagation, complaints about remembered content and divergence between local and cloud copies. Thresholds should trigger investigation, quarantine, correction of derivatives, suspension of synchronisation or other proportionate corrective action. Reporting follows the distinct legal tests in each regime; not every blocked event is a reportable incident.

Finally, fundamental-rights impact assessments under Article 27 and data-protection impact assessments where GDPR Article 35 applies should examine cumulative and cross-context effects. Cybersecurity controls can keep a record confidential and authentic while

[14] AI Act, Arts. 40–1; A Ojanen, "Technology Neutrality as a Way to Future-Proof Regulation: The Case of the Artificial Intelligence Act" (2025) 16 European Journal of Risk Regulation 1440.

https://doi.org/10.1017/err.2026.10159

its use remains discriminatory, irrelevant or unlawful. Integrated assessment is therefore necessary, but integration should preserve the different purposes of the AI Act, NIS2 and data-protection law.

VIII. Limits

The first limit is proportionality. Persistent memory does not make an otherwise low-risk application high-risk. Scrutiny should rise with the consequence of the task, the sensitivity and authority of remembered information, autonomy, tool permissions, cross-context reuse, difficulty of reversal and weakness of human review. This avoids the over-regulation that would follow from treating memory as a categorical trigger.

The second limit is institutional scope. NIS2 applies only to entities and sectors within its rules as transposed nationally, and sector-specific regimes such as the Digital Operational Resilience Act may alter the cybersecurity analysis. The present article therefore does not assert universal dual compliance. It argues that overlapping obligations must be analysed wherever both regimes apply.

The third limit is evidence. The cited security and misalignment studies include benchmarks, red-team simulations and recent preprints. They demonstrate plausible mechanisms and testing needs, not the frequency or severity of real-world harm. Regulatory implementation should remain evidence-based and should be updated through incident reports, conformity assessments, regulatory sandboxes and deployment studies.

The final limit is conceptual. Memory can mediate misalignment but is neither necessary nor sufficient for it. Emergent behaviour can arise without persistent storage, and well-designed memory can improve safety by preserving constraints, provenance and prior corrections. The regulatory objective is not to suppress memory. It is to make the assembled agent reliably governable in its actual context.

IX. Conclusion

Persistent agent memory is not training data merely because it influences a later prediction, and it is not a new risk category in the AI Act. It is an inference-time system feature that can make misalignment and emergent behaviour persistent, context-dependent and difficult to reconstruct. That framing preserves the article's concern while respecting the Act's context-specific architecture.

The AI Act already supplies substantial coverage. Article 9 captures foreseeable lifecycle risks; Articles 12–14 support evidence, deployer understanding and oversight; Article 15 clearly mandates robustness and cybersecurity; and Articles 25, 72 and 73 address value-chain change, monitoring and incidents. For entities within its scope, NIS2 adds management accountability, all-hazards security, access and asset controls, supply-chain governance, business continuity and incident reporting for the deployment environment.

The remaining task is implementation. Guidance, standards, documentation, contracts and monitoring should make memory provenance, retrieval scope, freshness, revocation and recovery testable in proportion to context. This approach neither exaggerates a regulatory vacuum nor assumes that general clauses are self-executing. It asks the more useful question: whether existing duties produce adequate controls and evidence when an agent's behaviour depends on what it remembers.

Financial support. This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

https://doi.org/10.1017/err.2026.10159

Competing interests. The authors declare none.

Ethics standards. Not applicable. The article is a doctrinal and conceptual legal analysis and does not involve human subjects, animal subjects or personal data.

Cite this article: A Adusei Brobbey, "Persistent Agent Memory under the EU AI Act and NIS2: Misalignment, Emergent Behaviour and Regulatory Fit". *European Journal of Risk Regulation*. https://doi.org/10.1017/err.2026.10159

https://doi.org/10.1017/err.2026.10159