Risk Assessment Report

1. Introduction

Effective risk management is a foundational element of modern governance, cybersecurity, and operational resilience. Regulatory frameworks such as the European Unionâ€™s NIS2 Directive require organizations operating critical or important services to systematically identify, assess, and manage cybersecurity risks that could affect the continuity and security of their services. Similarly, internationally recognized standards such as ISO/IEC 27001 and ISO/IEC 27005 emphasize structured risk management processes as a core component of an effective information security management system (ISMS). These frameworks require organizations not only to implement security controls but also to demonstrate that risks are understood, prioritized, and treated in a consistent and auditable manner.

This report presents the results of a structured risk assessment conducted using the Information Risk Assessment Methodology 2 (IRAM2). IRAM2 provides a systematic approach for analyzing information risks by evaluating potential threat events, their likelihood of occurrence, and their potential business impact across confidentiality, integrity, and availability dimensions. The methodology emphasizes traceability from system architecture and threat scenarios through to quantified risk levels and recommended treatments. By structuring the assessment around clearly defined models for threats, impacts, likelihood, and controls, IRAM2 enables transparent and repeatable risk analysis that supports informed decision-making.

Applying this methodology provides several benefits. It ensures that risks are evaluated in a consistent, evidence-based manner, supports alignment with regulatory and international standards, and allows stakeholders to understand how technical system characteristics translate into organizational risk exposure. The structured modeling approach also enables clearer prioritization of remediation activities, helping organizations focus resources on the most significant risks while documenting accepted risks and residual exposure.

The purpose of this report is therefore to present a clear and defensible overview of the systemâ€™s risk posture. It summarizes the system context, assessment methodology, threat and risk analysis results, and the recommended or accepted risk treatment actions. The report also documents residual risks and key assumptions that may influence the assessment. Together, these findings are intended to support governance decisions, guide remediation planning, and provide an auditable record of the organizationâ€™s risk management activities in alignment with regulatory and standards-based expectations.

2. Executive Context

This section captures the assessmentâ€™s administrative context and provides a concise narrative intended for executive stakeholders.

1. System name: `Riskonami`
2. System owner: `pieter@riskonami.com`
3. Assessment date: `2026-03-04`
4. System description: `Riskonami uses AI to risk assess systems and identify weaknesses.`

3. Methodology

This section explains the standards, scope boundaries, and confidence level so readers can interpret the findings appropriately.

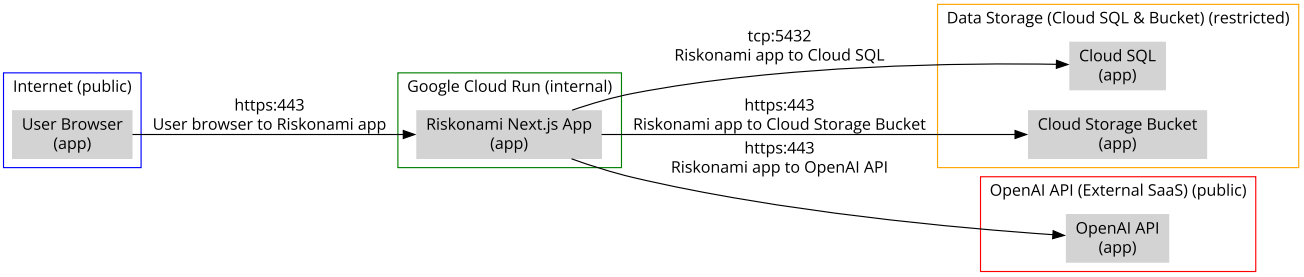
Methodology element	Detail
Standards and approach used	IRAM2-inspired, phase-based deterministic assessment workflow.
Phase coverage and analysis boundaries	Covers impact, threat modeling, likelihood, risk scoring, and treatment planning from session phase JSON.
Evidence quality and confidence narrative	Model evidence generated from structured phase outputs and locked phase artifacts.

JSON decoder	Status	Purpose
`impactmodel`	Decoded	System context and CIA/impact baseline
`threatmodel`	Decoded	Architecture, trust zones, and flows
`likelihoodmodel`	Decoded	Threat event likelihood scoring
`riskmodel`	Decoded	Risk synthesis and prioritization
`risktreatment`	Decoded	Treatment decisions and residual risk

4. System Overview

This section describes the system architecture and trust boundaries used as the basis for threat identification and risk analysis.

1. Architecture summary from `threatmodel.json`:
    1. System: `Riskonami`
    2. Trust zones: 4
    3. Flows: 4
    4. Description: `Riskonami uses AI to risk assess systems and identify weaknesses.`
2. Threat model diagram:
    1.



5. Findings Synthesis

This section summarizes the impact posture and relevant regulatory flags to frame the severity and constraints of the risk profile.

1. Impact assessment highlights from `impactmodel.json`:
    1. Confidentiality impact: `HIGH`
    2. Integrity impact: `HIGH`
    3. Availability impact: `MEDIUM`
    4. Final impact rating: `HIGH`
2. Legislative flags:
    1. GDPR: `Yes`

2. AI Act: `No`
3. NIS2: `Yes`
4. LED: `No`
5. SNISA: `No`
6. CRA: `No`

6. Threat and Risk Analysis

This section provides the quantified likelihood and risk synthesis that drives prioritization and treatment decisions.

1. Likelihood assessment table from `likelihoodmodel.json`:

Threat Event ID	Threat Event	TEL (Level/Score)	Calculated Likelihood	Adjusted Likelihood
ACC001	Account takeover via stolen or abused identities	Likely (4)	Very High (12)	Very High (12)
ACC002	Creation or use of rogue accounts	Possible (3)	Medium (6)	Medium (6)
ACC003	Privilege escalation within accounts	Likely (4)	High (8)	High (8)
ACC005	Abuse of legitimate access by authorised users	Likely (4)	High (8)	High (8)
ACC006	Session or token theft and reuse	Possible (3)	Medium (6)	Medium (6)
ACC007	Compromise of privileged administrative accounts	Likely (4)	High (8)	High (8)
ADV001	Injection attacks against the web application (XSS/DOM)	Likely (4)	High (8)	High (8)
ADV002	Credential stuffing and brute-force attacks against logins	Likely (4)	Very High (12)	Very High (12)
ADV003	Exploitation of authentication and session management flaws	Possible (3)	Medium (6)	Medium (6)
ADV004	Exploitation of authorization and access control flaws	Possible (3)	Medium (6)	Medium (6)
ADV005	Data exfiltration or tampering via legitimate application access	Likely (4)	Very High (12)	Very High (12)
ADV006	Exploitation of software supply chain and dependency weaknesses	Possible (3)	Medium (6)	Medium (6)
ADV007	Abuse or compromise of API keys and external service credentials	Likely (4)	High (8)	High (8)
ADV008	Inappropriate disclosure of sensitive data to third-party AI services	Likely (4)	Very High (12)	Very High (12)
ADV009	Exploitation of data validation and integrity weaknesses	Possible (3)	Medium (6)	Medium (6)
ADV010	Exploitation of configuration and security hardening weaknesses	Likely (4)	High (8)	High (8)

ADV011	Compromise via social engineering and phishing of staff	Likely (4)	Very High (12)	Very High (12)
ADV012	Compromise via malware or endpoint compromise	Possible (3)	Medium (6)	Medium (6)
ADV013	Exploitation of vulnerabilities in underlying infrastructure	Possible (3)	Medium (6)	Medium (6)
ADV014	Exploitation of logging and monitoring gaps	Possible (3)	Medium (6)	Medium (6)
ADV015	Data loss or corruption due to backup and recovery weaknesses	Possible (3)	Medium (6)	Medium (6)
ADV016	Denial of service against the web application and components	Likely (4)	High (8)	High (8)
ENV011	Cloud platform outage, quota exhaustion, or critical dependency failure	Possible (3)	Medium (6)	Medium (6)

2. Risk synthesis from `riskmodel.json`:

3. Risk: `Malicious or compromised insiders use their legitimate access or tokens to extract or alter sensitive data through the GenAI system.`

    1. Threat: `Data exfiltration or tampering via legitimate access`
    2. Impact: `Very High`
    3. Likelihood: `Almost Certain`
    4. Risk score: `20 (Critical)`

4. Risk: `Sensitive or regulated data is sent to third-party AI APIs or services without proper controls, potentially exposing it to those providers or their sub-processors.`

    1. Threat: `Inappropriate disclosure of sensitive data to third-party AI services`
    2. Impact: `Very High`
    3. Likelihood: `Almost Certain`
    4. Risk score: `20 (Critical)`

5. Risk: `Malicious or compromised insiders use their legitimate access or tokens to extract or alter sensitive data through the GenAI system.`

    1. Threat: `Data exfiltration or tampering via legitimate access`
    2. Impact: `Very High`
    3. Likelihood: `Likely`
    4. Risk score: `16 (Critical)`

6. Risk: `Sensitive or regulated data is sent to third-party AI APIs or services without proper controls, potentially exposing it to those providers or their sub-processors.`

    1. Threat: `Inappropriate disclosure of sensitive data to third-party AI services`
    2. Impact: `Very High`
    3. Likelihood: `Likely`
    4. Risk score: `16 (Critical)`

7. Risk: `User or administrator accounts for the GenAI system are taken over using stolen passwords, tokens, or session hijacking.`

    1. Threat: Account takeover via stolen or abused identities
    2. Impact: High
    3. Likelihood: Almost Certain
    4. Risk score: 15 (Critical)

8. Risk: Attackers use automated attempts with lists of stolen or guessed credentials to gain unauthorized access to user or admin accounts in the GenAI system.

    1. Threat: Credential stuffing and brute-force attacks
    2. Impact: High
    3. Likelihood: Almost Certain
    4. Risk score: 15 (Critical)

9. Risk: Attackers use automated attempts with lists of stolen or guessed credentials to gain unauthorized access to user or admin accounts in the GenAI system.

    1. Threat: Credential stuffing and brute-force attacks
    2. Impact: High
    3. Likelihood: Likely
    4. Risk score: 12 (High)

10. Risk: Malicious or compromised insiders use their legitimate access or tokens to extract or alter sensitive data through the GenAI system.

    1. Threat: Data exfiltration or tampering via legitimate access
    2. Impact: High
    3. Likelihood: Likely
    4. Risk score: 12 (High)

11. Risk: Sensitive or regulated data is sent to third-party AI APIs or services without proper controls, potentially exposing it to those providers or their sub-processors.

    1. Threat: Inappropriate disclosure of sensitive data to third-party AI services
    2. Impact: High
    3. Likelihood: Likely
    4. Risk score: 12 (High)

12. Risk: Attackers trick staff through phishing, social engineering, or deepfakes to gain access to GenAI administration, data, or integration channels.

    1. Threat: Compromise via social engineering and phishing of staff
    2. Impact: High
    3. Likelihood: Likely
    4. Risk score: 12 (High)

13. Risk: Attackers trick staff through phishing, social engineering, or deepfakes to gain access to GenAI administration, data, or integration channels.

    1. Threat: Compromise via social engineering and phishing of staff
    2. Impact: High
    3. Likelihood: Possible
    4. Risk score: 9 (High)

14. Risk: Attackers use automated attempts with lists of stolen or guessed credentials to gain unauthorized access to user or admin accounts in the GenAI system.

    1. Threat: `Credential stuffing and brute-force attacks`
    2. Impact: `Medium`
    3. Likelihood: `Likely`
    4. Risk score: `8 (Medium)`

15. Risk: `Attackers trick staff through phishing, social engineering, or deepfakes to gain access to GenAI administration, data, or integration channels.`

    1. Threat: `Compromise via social engineering and phishing of staff`
    2. Impact: `Medium`
    3. Likelihood: `Likely`
    4. Risk score: `8 (Medium)`

16. Risk: `User or administrator accounts for the GenAI system are taken over using stolen passwords, tokens, or session hijacking.`

    1. Threat: `Account takeover via stolen or abused identities`
    2. Impact: `Medium`
    3. Likelihood: `Possible`
    4. Risk score: `6 (Medium)`

17. Risk: `User or administrator accounts for the GenAI system are taken over using stolen passwords, tokens, or session hijacking.`

    1. Threat: `Account takeover via stolen or abused identities`
    2. Impact: `Medium`
    3. Likelihood: `Possible`
    4. Risk score: `6 (Medium)`

7. Control Evaluation

This section documents recommended treatments and explicitly highlights risks accepted by decision-makers.

1. Recommended controls from `risktreatment.json`:

Risk ID	Risk description	Treatment decision	Recommended control	Control detail
`RISK-1`	Account takeover or credential abuse leading to unauthorized access to Riskonami and connected GenAI capabilities.	Mitigate	Enforced phishing-resistant MFA and conditional access	Roll out enforced multi-factor authentication (preferably phishing-resistant where possible) and conditional access policies for all administrative and high-privilege users accessing Riskonami and GenAI-integrated services.
`RISK-2`	Data exfiltration or tampering via legitimate but excessive or misconfigured access within Riskonami and GenAI-integrated workflows.	Mitigate	Fine-grained access control and data governance for Riskonami and GenAI data flows	Implement more granular role definitions, least-privilege access, and data governance policies for sensitive data used or referenced by Riskonami and any connected GenAI services.
`RISK-3`	Inappropriate disclosure of sensitive or regulated data to OpenAI or other external	Mitigate	GenAI usage policy, guardrails, and technical restrictions	Define and enforce a GenAI usage policy, implement request validation and redaction where feasible, and configure technical controls to prevent or minimize

	GenAI providers via prompts or training data.			sending regulated or highly confidential data to external GenAI providers.
RISK-4	Erroneous or biased GenAI outputs influencing risk decisions or documentation in Riskonami.	Mitigate	Mandatory human-in-the-loop review and documentation standards	Formalize procedures requiring human validation of GenAI-generated content for material risk decisions, and maintain clear documentation of human approval for key outputs.

2. Accepted risks:

Risk ID	Risk description
N/A	None explicitly accepted in available data.

8. Residual Risk Narrative

This section records the expected remaining risk after treatments and frames what must be monitored or escalated.

1. Residual risk outcomes:

Risk ID	Risk description	Residual score	Residual level	Treatment decision
RISK-1	Account takeover or credential abuse leading to unauthorized access to Riskonami and connected GenAI capabilities.	3	Medium	Mitigate
RISK-2	Data exfiltration or tampering via legitimate but excessive or misconfigured access within Riskonami and GenAI-integrated workflows.	3	Medium	Mitigate
RISK-3	Inappropriate disclosure of sensitive or regulated data to OpenAI or other external GenAI providers via prompts or training data.	3	Medium	Mitigate
RISK-4	Erroneous or biased GenAI outputs influencing risk decisions or documentation in Riskonami.	3	Medium	Mitigate

9. Remediation Roadmap and Next Steps

This section translates risk treatments into an actionable plan, clarifying sequencing, ownership, and validation checkpoints.

1. Remediation roadmap tasks:

Priority	Task	Why this task	Risk(s) addressed	Owner	Target timeframe	Status	Dependency	Validation checkpoint
1	Enforced phishing-resistant MFA and conditional access: Roll out enforced multi-factor	Reduces risk exposure documented in treatment model.	RISK-1 - Account takeover or credential abuse leading to unauthorized access to	IAM Lead	TBD	Planned	None	Recalculate likelihood and residual score after implementation.

	authentication (preferably phishing-resistant where possible) and conditional access policies for all administrative and high-privilege users accessing Riskonami and GenAI-integrated services.		Riskonami and connected GenAI capabilities.					
2	Fine-grained access control and data governance for Riskonami and GenAI data flows: Implement more granular role definitions, least-privilege access, and data governance policies for sensitive data used or referenced by Riskonami and any connected GenAI services.	Reduces risk exposure documented in treatment model.	RISK-2 - Data exfiltration or tampering via legitimate but excessive or misconfigured access within Riskonami and GenAI-integrated workflows.	IAM Lead	TBD	Planned	After higher-priority roadmap tasks.	Recalculate likelihood and residual score after implementation.
3	GenAI usage policy, guardrails, and technical restrictions: Define and enforce a GenAI usage policy, implement request validation and redaction	Reduces risk exposure documented in treatment model.	RISK-3 - Inappropriate disclosure of sensitive or regulated data to OpenAI or other external GenAI providers via prompts or training data.	Data Protection Officer	TBD	Planned	After higher-priority roadmap tasks.	Recalculate likelihood and residual score after implementation.

	where feasible, and configure technical controls to prevent or minimize sending regulated or highly confidential data to external GenAI providers.							
4	Mandatory human-in-the-loop review and documentation standards: Formalize procedures requiring human validation of GenAI-generated content for material risk decisions, and maintain clear documentation of human approval for key outputs.	Reduces risk exposure documented in treatment model.	RISK-4 - Erroneous or biased GenAI outputs influencing risk decisions or documentation in Riskonami.	Data Protection Officer	TBD	Planned	After higher-priority roadmap tasks.	Recalculate likelihood and residual score after implementation.

10. Conclusion

This section provides decision-support wrap-up: what the risk posture is, what should be decided now, and how often to reassess.

1. Final risk posture summary: `Overall impact posture: HIGH. Top risks: ADV005_1 (20/Critical), ADV008_1 (20/Critical), ADV005_3 (16/Critical).`
2. Decision support statements for stakeholders: `Prioritize controls for highest-scoring risks and explicitly track accepted-risk rationale.`
3. Recommended governance cadence for reassessment: `Quarterly reassessment, with immediate reassessment after major architecture or control changes.`

11. Appendix A: Assumptions and Limitations

This appendix records uncertainty and limitations so stakeholders understand where conclusions may change with better evidence.

1. Explicit assumptions made during analysis: `None explicitly captured.`
2. Known data gaps and constraints: `Some sections depend on phase completeness and locked JSON outputs.`
3. Impact of assumptions on risk confidence: `Confidence is reduced when upstream phase evidence is missing or stale.`