System and Scope

Overview	Scope	Inputs	Outputs	Ownership	Supervisor
LLM-Powered Incident response Assistant for SMEs: Ingests network logs from IDS, passes it to LLM for summarisation, Outputs into a DB, Displays on a Web Dashboard.	Local Deployment on VM for testing., evaluates how LLMs can assist SMEs during incident response.	JSON logs	LLM-generated dashboard data.	Matthew Fish	Nathan Clarke

Data Flow

Source	Type	Flow	PII Risk	Notes
Network logs	Structured JSON	JSON Network logs to LLM	Low–Medium	IP addresses may indirectly identify individuals.
LLM	JSON input/ Text output	Receives JSON, Returns summarised data	Low	Model runs locally so no external data transfer.
Database	Structured	Stores processes data	Medium	Must ensure encryption and restricted access.
Web dashboard	Structured	Displays processed data	Low	User authentication needed to prevent exposure.
Test VMs	Unstructured	Simulated attacks	None	Used for testing only

Potential Risk Categories

Category	Description
Bias & Fairness	LLM may dramatise risk scores.
Privacy Violations	Logs could contain identifiable IPs or metadata.
Explainability	LLM output may be unclear or untraceable to specific log evidence.
Security	Exposure of database, API endpoints, or LLM prompt injection, User tampering.
Regulatory / Compliance	Data handling must align with GDPR.
Operational / Reliability	LLM, IDS, API or Dashboard downtime may stop attacks from being prevented.
Data Integrity	Corrupted or manipulated logs could cause false insights.
Supply Chain	Third party risks (docker images, pre-trained LLM)
People	Lack of user training and lack of incident response plan for if the tool got compromised

Likelihood and Impact					Rules	
Risk	**Likelihood**	**Impact**	**Overall Risk**	**Rationale**	**1 to 10**	**Low, Medium, High, Significant**
Model bias in classification	5	7	High	Depends on the data the LLM has been trained with. Will use a RAG to reduce bias as much as possible.		
Privacy leak via IP data	4	8	High	Rare but could have a serious impact. Will try to anonymise IPs where possible.		
Model hallucination / misinterpretation	5	7	High	Will implement a way to check the processed data.		
Database compromise	2	10	Medium	Huge affect on the company as an attacker could do a ransomware attack or use the logs for other attacks.		
LLM prompt injection or misuse	2	7	Medium	Stored locally so very unlikely and an implemented check on the data would help with security and outputs.		
Unauthorised access to dashboard, llm or IDS	2	10	Medium/ High	Will all be deployed locally and auth access will be required for dashboard. Access controls would be implemented for only admin to access llm or IDS.		
Data loss / corruption	2	5	Medium	Could use backups of data but requires further storage and security.		
System downtime	5	8	High	Downtime for updates, internet loss, power loss affects the whole system.		

Non-compliance with GDPR and computer misuse for users	2	10	Medium	Implment secure handling of data, access controls and systems in place to prevent tool being misused.
Supply chain risks	4	6	High	Dependencies on third-party vendors or libraries could fail or be compromised, affecting system stability and security.
Insufficient Logging	3	6	Medium	Limited logging makes it harder to detect, investigate, or respond to issues
User Error or misconfig	4	5	Medium	Misconfigurations or mistakes by users can cause outages or security issues
Ethical Misuse of tools	2	8	Medium	Tools could be used unethically (e.g., privacy violations)

Existing Mitigation Control

Control	Description	Effectiveness	Enforced?	Works as intended?
Docker container isolation	Each service runs in its own container.	High	To be configured	To be Verified
Everything run locally	No public exposure to data.	High	To be configured	To be Verified
Basic auth on API/dashboard	Restricts access.	Medium	To be configured	To be Tested
Access Control	RBAC for system access. Principle of least pivilege.	Medium	To be configured	To be Tested
Pseudonymisation of IPs	Truncate or hash IP addresses before storing.	Medium	To be configured	To be Tested
Database Security	Least-privilege accounts , encryption at rest	Medium	To be configured	To be tested
System updates	Patch Docker images regularly, IDS, LLM, Python.	High	Ongoing	Yes
Input Sanitisation and validation	Validate and sanitise all logs before entering LLM.	Medium	To be configured	To be verified
Logging and monitoring	System healthchecks and logs for user actions.	Medium	To be configured	To be Verified
Backup and Recovery	Regular backups of database and critical files	High	To be configured	To be Verified

Risk Register

Risk Description	Likelihood	Impact	Overall Risk	Controls	Owner	Mitigation / Plan	Timeline
LLM produces biased or inconsistent	5	7	High	Will review it manually.	Matthew	Validate LLM outputs vs. known threats. RAG	By Feb
IP data exposes personal info	4	8	High	IP anonymisation, stored locally.	Matthew	Hash Ips	Jan
Unauthorised LLM, IDS, dashboard, db access	2	10	High	Auth / Access Controls,docker isolation	Matthew	Implement JWT/API-key login. Access controls	Jan
Prompt injection or malicious input	2	7	Medium	Sanitise model inputs.	Matthew	Validate all logs before / after LLM processing	Jan
Data loss or	2	5	Medium	Backups	Matthew	Backups	Feb
Downtime or service failure	5	8	High	Monitoring, healthchecks	Matthew	Monitor uptime, add healthchecks, updates scheduled at low volume times.	Feb
Non-compliance with GDPR / Computer Misuse	2	10	Medium	Ethical review, documentatio	Matthew	Ethical review, Document compliance steps.	Ongoing
Supply chain risks	4	6	High	System updates, vulnerability scanning	Matthew	Trusted docker images, regular scans.	Ongoing
Insufficient Logging	3	6	Medium	Audit Logging	Matthew	Limited logging makes it harder to detect, investigate, or respond to issues. Access to tools are logged	Ongoing
User Error or misconfig	4	5	Medium	Training and documentation	Matthew	Clear documentation for setup and use. Configuration to avoid manual	Ongoing
Ethical Misuse of tools	2	8	Medium	Access controls, policies	Matthew	Restrict tool access to authorised users. Define and communicate acceptable use	Ongoing