AI National Security
Risk Assessment Framework

Type: Assessment Framework | **Audience:** Security Teams, Defense, Critical Infrastructure

1. Dual-Use Risk Evaluation

- ☐ Could this AI capability be weaponized?
- ☐ Does it enhance surveillance capabilities?
- ☐ Could it disrupt critical infrastructure?
- ☐ Does it provide strategic military advantage?
- ☐ Are there export control implications (ITAR, EAR)?

2. Adversarial Threat Assessment

Threat Type	Description	Risk Level
Model Poisoning	Compromised training data introducing backdoors	[ ]High [ ]Med [ ]Low
Adversarial Attacks	Inputs designed to fool AI systems	[ ]High [ ]Med [ ]Low
Model Theft	IP exfiltration of model weights/architecture	[ ]High [ ]Med [ ]Low
System Manipulation	Backdoors, trojans, unauthorized access	[ ]High [ ]Med [ ]Low
Supply Chain Attack	Compromised dependencies or components	[ ]High [ ]Med [ ]Low

3. Supply Chain Vulnerability Checklist

Training Data Sources:

- ☐ Data sources documented and verified
- ☐ Foreign data dependencies assessed
- ☐ Data integrity verification in place

Compute Infrastructure:

- ☐ Cloud provider jurisdiction verified
- ☐ Foreign ownership/control assessed (CFIUS)
- ☐ Security certifications verified (FedRAMP, etc.)

Model Components:

- ☐ Third-party libraries audited
- ☐ Open-source dependencies reviewed

4. Risk Scoring Matrix

Dimension	Description	Score (1-5)
Sensitivity	How critical is the capability? (5=critical national security)	[ ]
Vulnerability	How exposed is it to threats? (5=highly vulnerable)	[ ]
Impact	What is consequence of compromise? (5=catastrophic)	[ ]

Overall Risk = Sensitivity × Vulnerability × Impact = _____

Score Range	Risk Level	Required Controls
1-25	Low	Standard security controls
26-75	Moderate	Enhanced controls + monitoring
76-125	High	Specialized controls + continuous monitoring

5. Classification Guidance

Classification	Criteria	Handling
Classified	Used for defense/intelligence; processes classified data	Cleared personnel only; secure facilities
Sensitive	Could threaten national security if compromised	Need-to-know; access controls
CUI	Dual-use potential; controlled unclassified	Marking required; handling protocols
Public	General commercial application	Standard business practices

AI System: _____ Classification: _____

Assessed By: _____ Date: _____