ACADEMIC PAPER OVERVIEW • CYBERSECURITY • PRIVACY

AI Risk Management Frameworks for Data Privacy

A comparative synthesis across international regulations, standards, technical resources and consulting frameworks.

Liviu-Mihai Popescu & Remus Brad • Universitatea Lucian Blaga din Sibiu, Facultatea de Inginerie
Liviumihai.Popescu@ulbsibiu.ro

PROBLEM STATEMENT

Proliferation of AI Risk Frameworks

The Governance Challenge

Rapid AI adoption outpaces traditional cybersecurity models. Organizations face overlapping, fragmented guidance across multiple regulatory jurisdictions.

Regulatory Pressure

Mandatory conformity audits replace voluntary ethical charters.

Privacy Complexity

Model weights and inference APIs create unique data leakage vectors.

METHODOLOGY

Selection Criteria for Analysis

Global Influence

Worldwide institutional reach.

AI Specificity

Targeted machine learning risks.

Lifecycle Span

Coverage from design to retirement.

Audit Usability

Verifiable engineering controls.

TIER 1 BASELINES

Regulations and Consensus Standards

EU AI Act (2024)

Binding Regulation: Mandatory rules for European deployments.

Risk Tiers: Strict constraints on high-risk AI.

GDPR Alignment: Enforceable technical data governance.

NIST AI RMF 1.0

Voluntary Model: Widely used global benchmark.

Core Functions: Govern, Map, Measure, Manage.

Privacy Profile: Integrates with NIST Privacy Framework.

ISO/IEC 42001

Certifiable AIMS: Verifiable management systems standard.

Continuous Audit: Structured operational risk controls.

Lineage Verification: Rigorous lifecycle data documentation.

COMPARATIVE MATRIX

Regulatory & Standard Comparison

Framework	Legal Nature	Core Pillars	Privacy Controls
EU AI Act	**Binding Legal Mandate**	4 Risk Tiers	GDPR extension with technical safeguards
NIST AI RMF	**Voluntary Guidance**	Govern, Map, Measure, Manage	Privacy-enhanced trustworthy AI criteria
ENISA FAICP	**Voluntary Guidance**	Foundations, Specific, Sectoral	Federated learning and extraction defense
ISO/IEC 42001	**Certifiable Standard**	AIMS Process Requirements	Data governance and impact assessment

TIER 2 ARCHITECTURE

Technical Resources & Cloud Frameworks

MITRE ATLAS

Adversarial tactics knowledge base mapping model exfiltration threats.

Google SAIF

Six architectural pillars extending Zero Trust controls to ML workflows.

AWS CAF-AI

Operational blueprint embedding data privacy within cloud landing zones.

TIER 3 PRACTICE

Industry Professional Perspectives

Accenture & EY

Model reconstruction mitigation and board-level risk controls.

IBM Watsonx

Automated model lineage tracking and runtime data masking.

Palo Alto Networks

Runtime inference inspection and training pipeline defense.

CRITICAL FINDING

The Operational Implementation Gap

GAP

Statutory Principles vs. Engineering Realities

Regulations mandate data minimization and quality, but lack concrete architectural designs for model pipelines and runtime defenses.

THREAT TAXONOMY

Three Distinct AI Privacy Vectors

VECTOR 01

Model Inversion

Extraction of private training records via targeted prompts and mathematical inversion.

VECTOR 02

Machine Unlearning

Difficulty of verifiable data deletion from trained weights under GDPR erasure mandates.

VECTOR 03

Tool-Call Exfiltration

Context manipulation via indirect prompt injections leaking user data through external APIs.

ARCHITECTURE EVOLUTION

Composable Systems & Dynamic Trust

Distributed Topologies

AI workflows shift from monoliths to modular autonomous agents and vector databases.

Dynamic Trust Boundaries

Agents dynamically negotiate context sharing based on verified target exposure levels.

RECOMMENDATIONS

Core Pillars for AI Privacy Defense

01

Pipeline Isolation

Hardware-enforced confidential computing boundaries.

02

Runtime Guardrails

Active filtering of agent tool-calling outputs.

03

Automated Threat Models

Continuous evaluation embedded in workflows.

CONCLUSION

From Periodic Audits to Continuous Dynamic Trust

Privacy in modern AI cannot rely solely on static checklists. It requires dynamic runtime defense across composable agentic systems.

Thank You • Questions & Discussion

Liviu-Mihai Popescu & Remus Brad • Universitatea Lucian Blaga din Sibiu, Facultatea de Inginerie
Liviumihai.Popescu@ulbsibiu.ro