Liability Quant

Methodology Explainer

Quantification of Data Privacy Risk

Version 1.0 — Confidential

Overview

Liability Quant (LQ) provides an institutional due diligence standard for quantifying liability arising from personal data.

Use Cases

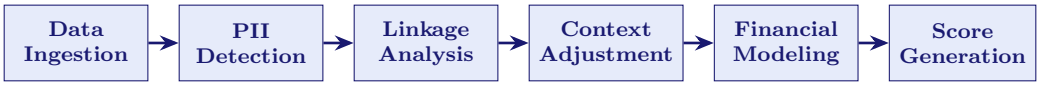
- M&A due diligence
- Acquisition of proprietary data or LLMs
- Cyber insurance underwriting
- General compliance assessments

Outputs

- **LQ Score** (0–100): Higher = lower risk
- **Financial Liability**: Dollar exposure estimate

The LQ Score is a directional risk indicator, analogous to credit ratings—consistent, reproducible, methodology-driven.

Risk Assessment Pipeline



Each stage is traceable. Identical inputs produce identical outputs.

Detection Framework

LQ employs a **two-pass, four-tier detection hierarchy** to identify **70+** PII types and **25+** linkage combinations. Deterministic; no LLMs.

First Pass: Atomic PII Detection

1. Structural	Headers for metadata and schema mapping
2. Pattern	Deterministic regex—accounts, IDs, phones
3. Entity	NLP-based named entity recognition
4. Content	Keyword discovery and partial matching

Second Pass

- Detect **linkage combinations**
- Conduct **subject inference**

Validation Layer

Checksums (Luhn, Modulo-97, Modulus-11) reduce false positives by **10–15%**.

The Mosaic Effect

Seemingly innocuous fields, when combined, uniquely identify individuals:

- **87%** of U.S. population identifiable via ZIP + DOB + gender
- **95%** identifiable from four location-time points

LQ identifies **toxic pairs** and applies **multiplicative amplification**—linkable combinations transform the attack surface globally, not incrementally.

Research Citations

Sweeney (1997)
de Montjoye et al. (2013)
Gymrek et al. (2013)

Multiplicative Risk Model

Risk is calculated as a **product of compounding factors** rather than a linear sum—accounting for the Mosaic Effect.

$$R_{\text{total}} = R_{\text{base}} \times A_{\text{linkage}} \times M_{\text{context}}$$

R_{base}

Weighted aggregate of detected PII anchored to statutory fines (GDPR, HIPAA, CCPA). Health/biometric data weighted heavily; quasi-identifiers lightly.

A_{linkage}

Mosaic Effect amplifier for re-identification risk from linkable clusters like DOB + ZIP + gender.

M_{context}

Product of operational multipliers: jurisdiction, data subject risk, and use case.

Rating Classification

LQ Score	Risk Class	Transactional Action
90–100	Minimal	Standard governance
70–89	Low	Standard Representations & Warranties
50–69	Moderate	Enhanced controls and monitoring
30–49	High	**Indemnity / Escrow Required**
0–29	Critical	**Transactional Remediation**

Illustrative Example: Customer database with SSNs, emails, DOBs, timestamps, and SSN+DOB toxic pair. EU deployment, 60% high-risk subjects. **Result:** LQ Score 38 (High Risk), Est. Liability $8.2M.

Transactional Application

For critical-risk datasets, **Transactional Remediation** options include:

Protections (preserve full data utility)

- Technical hardening (tokenization, encryption-at-rest, access controls)

Legal/Contractual:

- Specific indemnities
- Escrow arrangements
- Compliance warranties
- Remediation timelines

Data Modification (privacy-utility trade-off)

- Suppression
- Generalization
- Differential privacy

These techniques reduce re-identification risk but may diminish analytical value of the dataset.

Scope & Limitations

What LQ Analyzes

- Structured datasets and database exports (CSV, Excel, SQL)

Point-in-Time

- Reflects state at analysis
- Periodic reassessment recommended

Limitations

- Actual penalties depend on enforcement discretion
- Estimates anchored to regulatory frameworks
- Cannot detect encoded, encrypted, or obfuscated data
- Cannot verify sample is representative or complete

Liability Quant

Data Privacy Risk Quantification

This methodology provides quantitative risk indicators for due diligence purposes and does not constitute legal advice. Consult qualified counsel for jurisdiction-specific guidance.