THREAT MODELING CONNECT | POWERED BY THREATMODELER + IRIUSRISK

Viele Threats. Ein Budget. Was jetzt?

Quantitatives Risiko im Threat Modeling —
mit Monte Carlo, Loss Exceedance Curves und Risquanter Register

Daniel Agota · TM-Community-Session · 14.07.2026
danago@risquanter.com

Der doppelte Schmerz

Das Threat Model ist fertig. 40 Threats.

Schmerz 1 — Und jetzt?

- Welcher Threat zuerst? (Ranking)
- Ist A schlimmer als B? (Vergleich)
- Was tragen alle zusammen bei? (Aggregation)

Der doppelte Schmerz

Schmerz 2 — Und was sagen wir dem Management?

- Threat Modeling kostet Zeit und Geld
- „Was hat das TM-Programm gebracht?“
- Ohne Antwort: kein Budget für Programm, Tooling, Maßnahmen

Kein theoretisches Problem

Reale TM-Ergebnisse: lange Listen, heterogene Threats

Information Disclosure neben DoS neben Spoofing — Äpfel und Birnen

„Risiko“ ist in der TM-Community ein hitzig diskutiertes Thema (vgl. Keynote von Adam Shostack)

THREAT MODELING CONNECT

Ein Threat ist noch kein Risiko

Neue These aus der TM-Community: „Threat Modeling erfüllt den risikobasierten Ansatz (CRA, NIS2, DORA) per Definition.“

Halb richtig — und genau deshalb gefährlich:

- Threat = ein Szenario. „Was kann schiefgehen?“ (Spoofing, Tampering, DoS …)
- Risiko = Szenario + Eintrittswahrscheinlichkeit + Schadenshöhe, auf EINER Skala

Warum Risikomatrizen — und wo sie kippen

Die Idee: Wahrscheinlichkeit × Schaden in wenige Stufen (L/M/H) — ein Bild, das jeder auf einen Blick versteht.

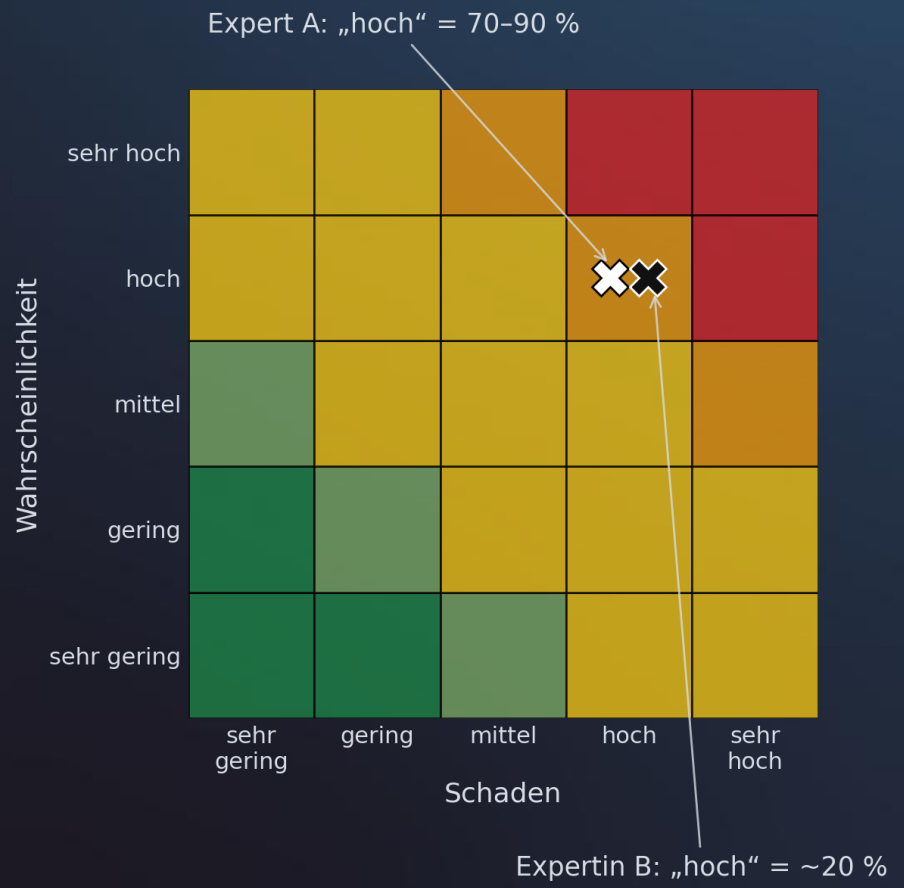
Zwei eingebaute Probleme:

- „L/M/H“ ist kein gemeinsamer Maßstab — dieselben Worte, verschiedene Welten
- 3–5 Stufen für alles von Bagatelle bis Katastrophe — zu grob zum Ranken

L/M/H — kein gemeinsamer Maßstab

Zwei Experten, dieselbe Matrix, verschiedene Welten:

- Expert A: „hohe Wahrscheinlichkeit“ = 70–90 %
- Expertin B (risikoavers): beim PII-Breach sind 20 % pro Jahr bereits das obere Ende der Skala
- **Beide kreuzen „High“ an — die Matrix versteckt den Dissens**



Warum CVSS — der Wunsch nach EINER Zahl

Warum es alle nutzen: standardisiert, vendor-neutral, eine Zahl von 0–10, die in SLAs und Patch-Fenster passt.

Verlockend: „7.5 — kritisch, diese Woche patchen.“ Vergleichbar, ohne Diskussion.

CVSS unter der Lupe

$$BaseScore = (0.6 \cdot Impact + 0.4 \cdot Exploitability - 1.5) \cdot f(Impact)$$

$$Impact = 10.41 \cdot (1 - (1-C)(1-I)(1-A))$$

$$Exploitability = 20 \cdot AccessComplexity \cdot Authentication \cdot AccessVector$$

$$f(Impact) = 0 \text{ falls } Impact = 0, \text{ sonst } 1.176$$

Woher kommen 10.41, 1.176, 0.704? Fest verdrahtete Gewichte — nicht herleitbar.

CVSS unter der Lupe

Eingabe: ein paar Kategorien ankreuzen

- z. B. AccessComplexity: high / medium / low → 0.35 / 0.61 / 0.71

Formel mit festen Gewichten

Ausgabe: 0.0–10.0 in Zehntelschritten — endlich viele Eimer

Kategorien rein, Kategorien raus — die Dezimalstellen sind Kosmetik

Der Score ordnet nur — er misst nicht

Ein Score ist „ein Algorithmus, der kategorisierte Urteile auf eine Standardskala abbildet“ (Cox 2009) — 101 Eimer in fester Reihenfolge

Die Reihenfolge stimmt — aber es gibt kein Maß darüber:

- **4 ist nicht halb so schlimm wie 8**
- **der Schritt 3 → 4 wiegt nicht so viel wie 7 → 8**

Rang ja, Verhältnisse und Abstände nein — nicht mittelbar, nicht addierbar

FIRST selbst: CVSS misst Severity, nicht Risk (FIRST 2019)

„Not justified, either formally or empirically“ (Spring et al. 2021)

Und selbst wenn Scores etwas messen würden …

Ranking nach Scores kann schlechter abschneiden als ZUFÄLLIGE Auswahl der Maßnahmen (Cox 2009)

Warum? Die Liste ist blind für Zusammenhänge:

- EIN Fix entschärft drei Threats — die Top-10-Liste sieht es nicht
- fünf „Mediums“ teilen dieselbe Ursache und feuern gemeinsam — die Liste zählt sie einzeln

⇒ Zum Vergleichen, Aggregieren und Ranken braucht es ein echtes Maß — Geld funktioniert in vielen Domänen.

Was brauchen wir also?

Anforderungen:

vergleichen · aggregieren · ranken · Szenarien durchspielen

Monte Carlo Simulation:

die Standardmethode, die genau das liefert

Eintrittspreis pro Risiko — nur zwei Dinge:

- Eintrittswahrscheinlichkeit (pro Jahr)
- eine Schadensverteilung, aus der gesampelt wird

So rechnet Monte Carlo — ein Münzwurf pro Jahr

Ein Münzwurf pro simuliertem Jahr:

- Kopf → einen Schaden aus der Verteilung ziehen
- Zahl → 0

Das mal tausende „Jahre“ — dann zählen, was zusammenkommt.

Nach Häufigkeit gruppiert ergeben die Werte ein Histogramm ≈ eine Verteilung.

Münzwurf pro Jahr

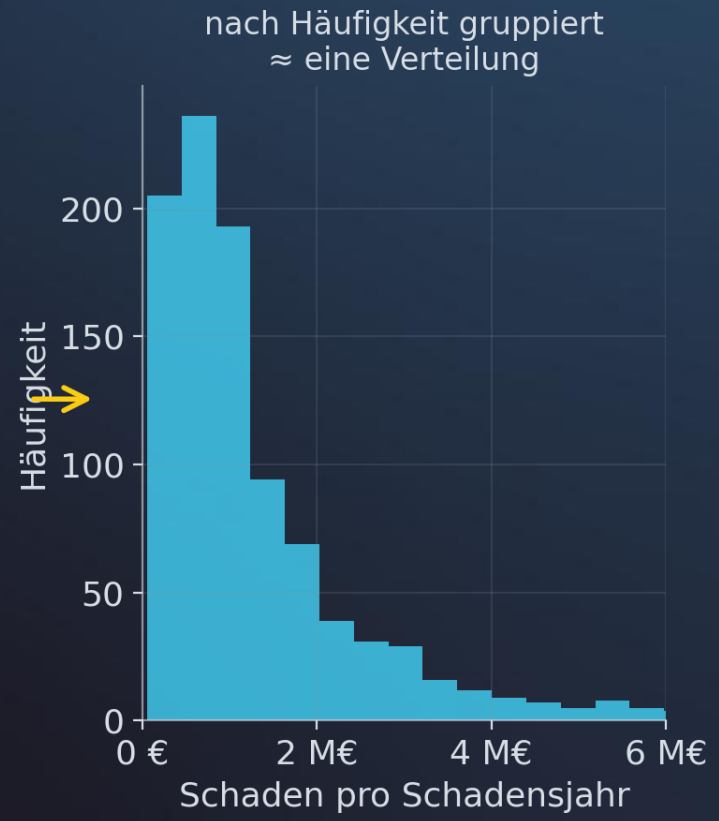
P = 10 %

Kopf → Schaden ziehen

Zahl → 0

je Jahr EINE Zahl
(× 10 000 Jahre)

```
Jahr 1:   0
Jahr 2:   0
Jahr 3:   1,2 M€
Jahr 4:   0
Jahr 5:   0,4 M€
  …
```



Wer rechnet so? — ein reifes Feld

Monte Carlo für Verlustrisiken ist nichts Neues — wir borgen es aus Disziplinen, die seit Jahrzehnten so rechnen:

- Rück-/Versicherer: Katastrophenmodellierung (Sturm, Erdbeben, Pandemie)
- Banken: operationelles Risiko, Value-at-Risk
- Casinos: das Haus gewinnt, weil es die VERTEILUNG kennt — nicht den einzelnen Wurf

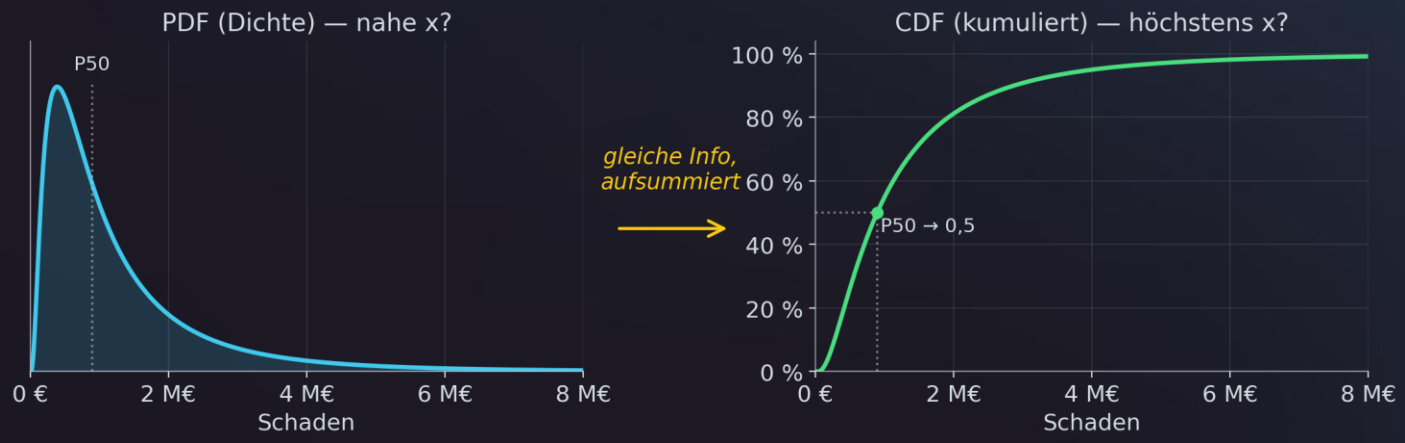
Neu ist nur, es sauber aufs Threat Modeling anzuwenden — mit denselben zwei Zahlen pro Risiko.

MC Input: Wahrscheinlichkeit + Schadenverteilung

Die Verteilung wird beschrieben:

- entweder mit der Dichte: „Wie häufig ist ein Schaden X?“
- oder kumulativ: „Wie wahrscheinlich ist HÖCHSTENS X?“

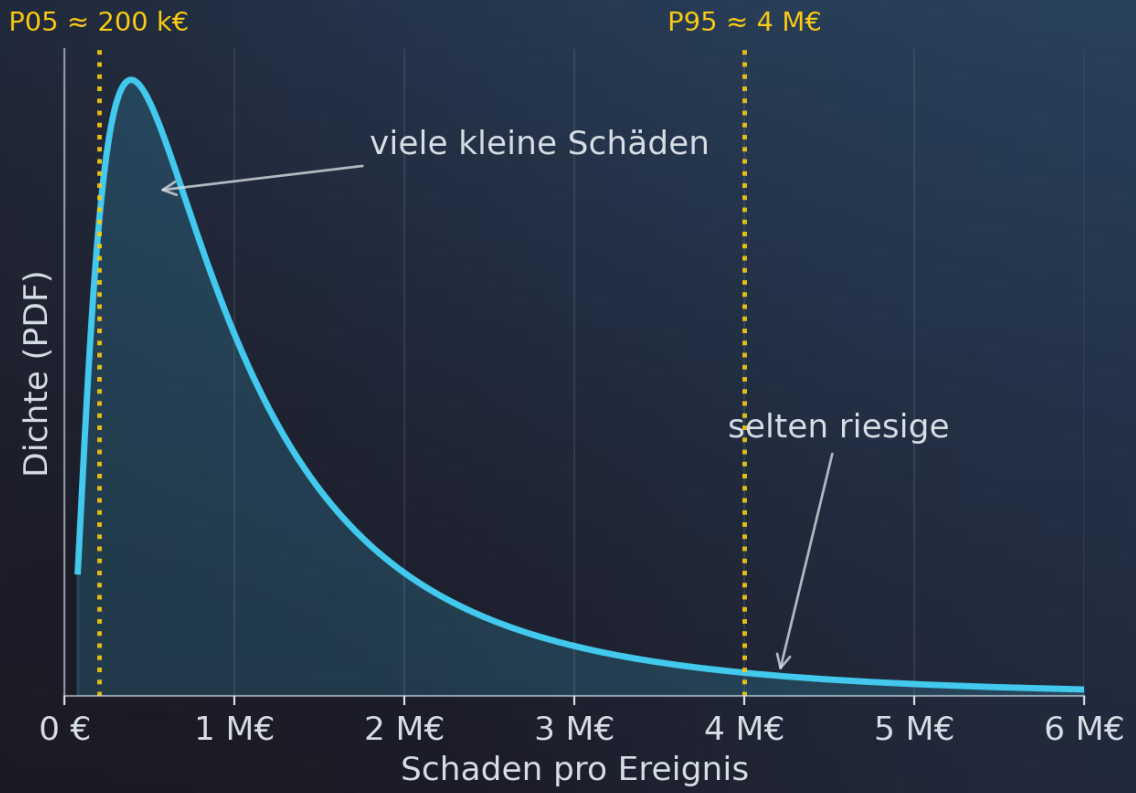
Beide tragen dieselbe Information — nur umgerechnet (die CDF ist die aufsummierte PDF).



Welche Verteilung? — Lognormal

Lognormal-Verteilung: die natürliche Form für Schäden

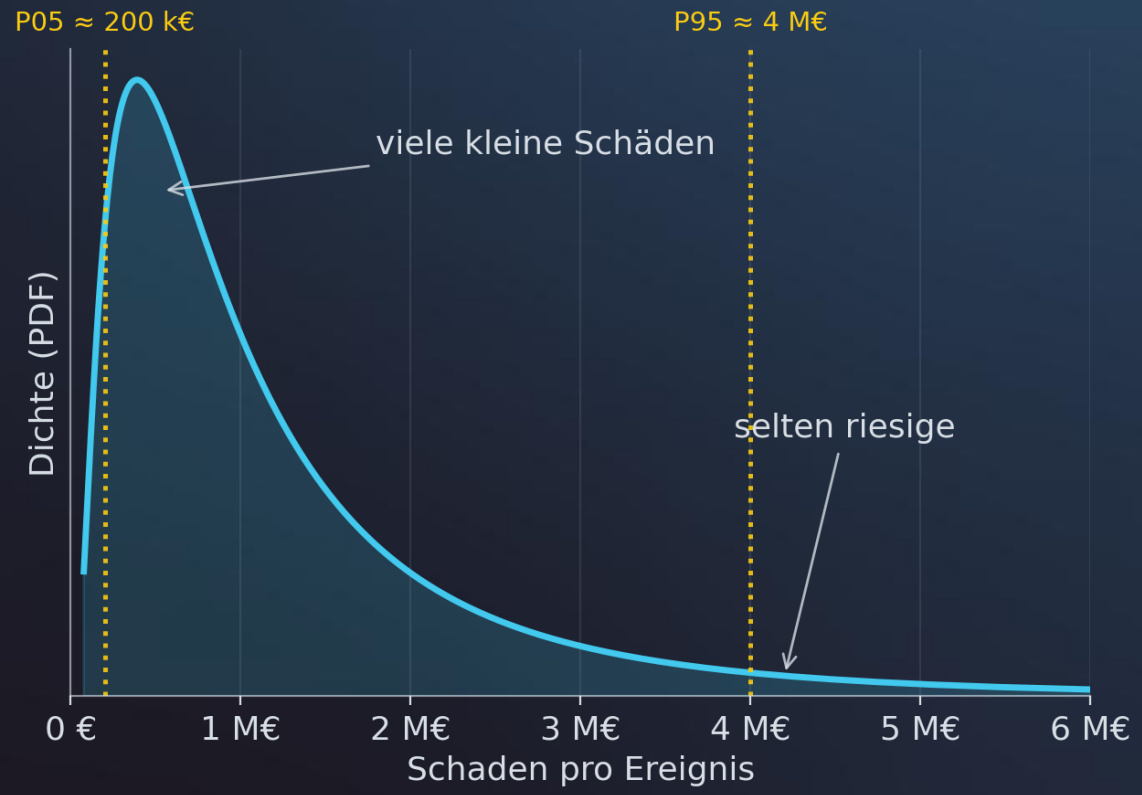
- nur positive Werte
- rechtsschief: viele kleine Schäden, selten riesige
- bewährter Standard der operationellen Risikopraxis



Welche Verteilung? — Lognormal

Zwei Aussagen genügen:

- „Passiert mit ca. 10 % pro Jahr.“ → Eintrittswahrscheinlichkeit
- „Ich bin zu 90 % sicher: falls es passiert, liegt der Schaden zwischen ~200 k€ und ~4 M€.“ → 90% Confidence Interval (P05–P95) der Lognormal-Verteilung

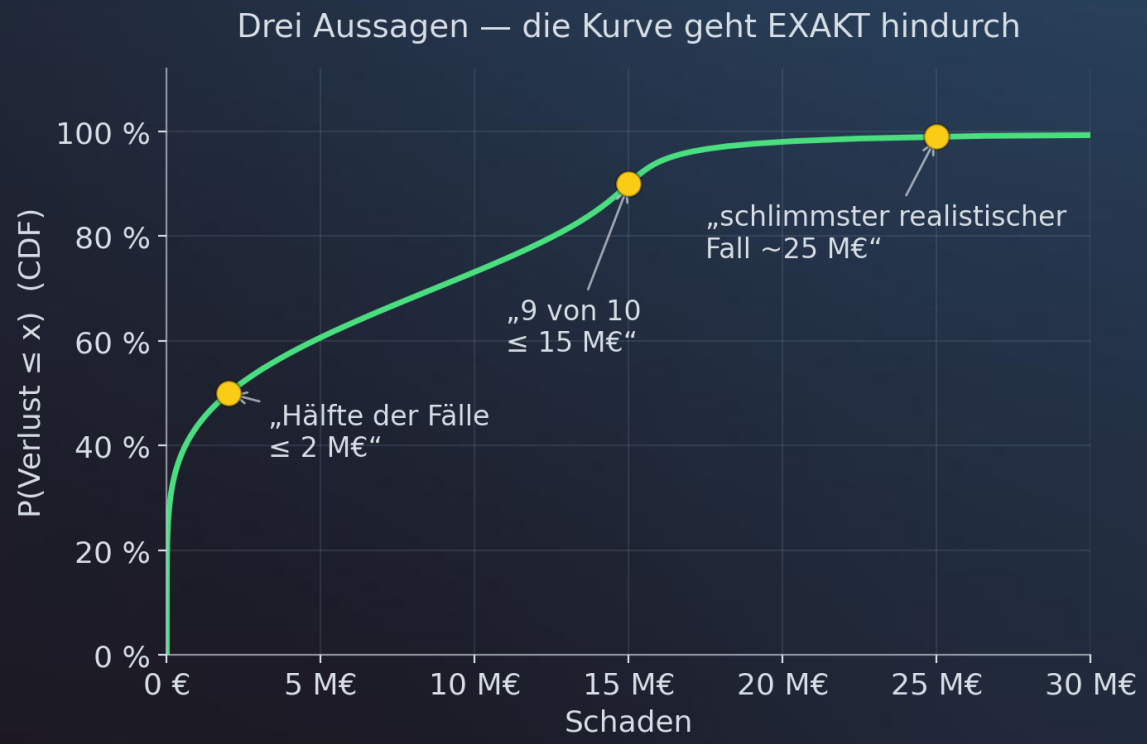


Welche Verteilung? — Metalog

Manchmal klingt Expertenwissen anders:

- „In der Hälfte der Fälle bleibt es unter 2 M€.“
- „In 9 von 10 Fällen unter 15 M€.“
- „Schlimmster realistischer Fall: ~25 M€.“

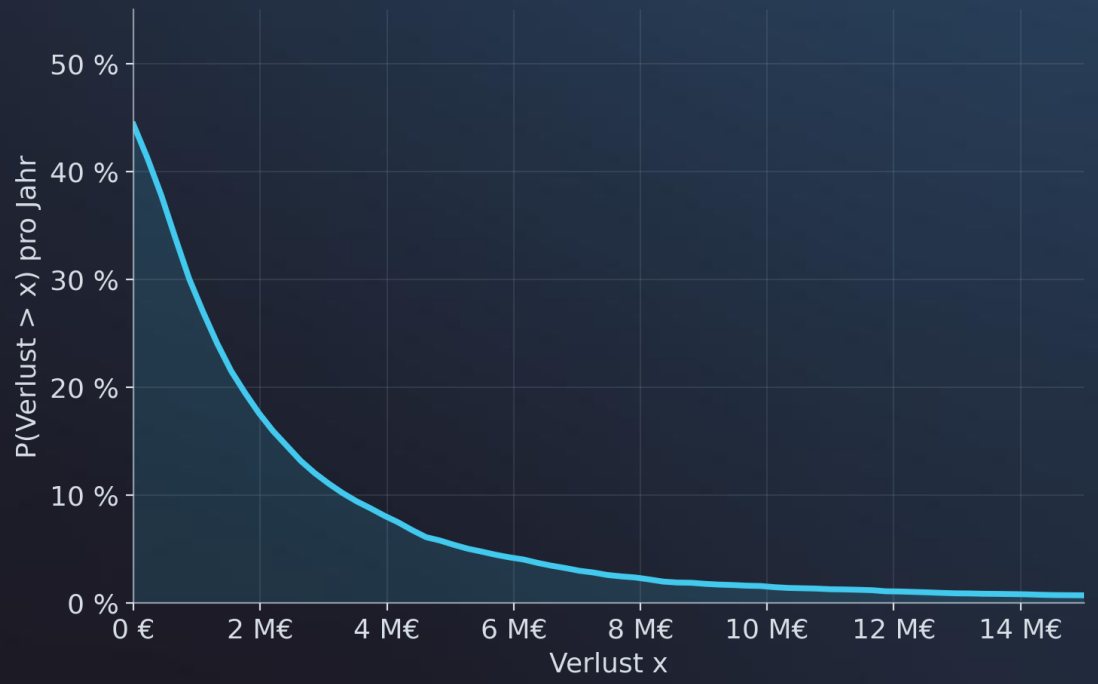
Die Metalog-Verteilung (Keelin 2016) legt die CDF-Kurve EXAKT durch die eigenen Aussagen — Verteilungstyp nicht explizit gegeben



MC Output: Die Loss Exceedance Curve (LEC)

Es wird für jeden Schadenslevel x gezählt, in wie vielen der 10.000 Jahre der Schaden über x lag.

- 760 von 10.000 Jahren über 20 M€ → 7,6 % Chance pro Jahr
- x-Achse: Schadenshöhe
- y-Achse: jährliche Wahrscheinlichkeit, MEHR als x zu verlieren

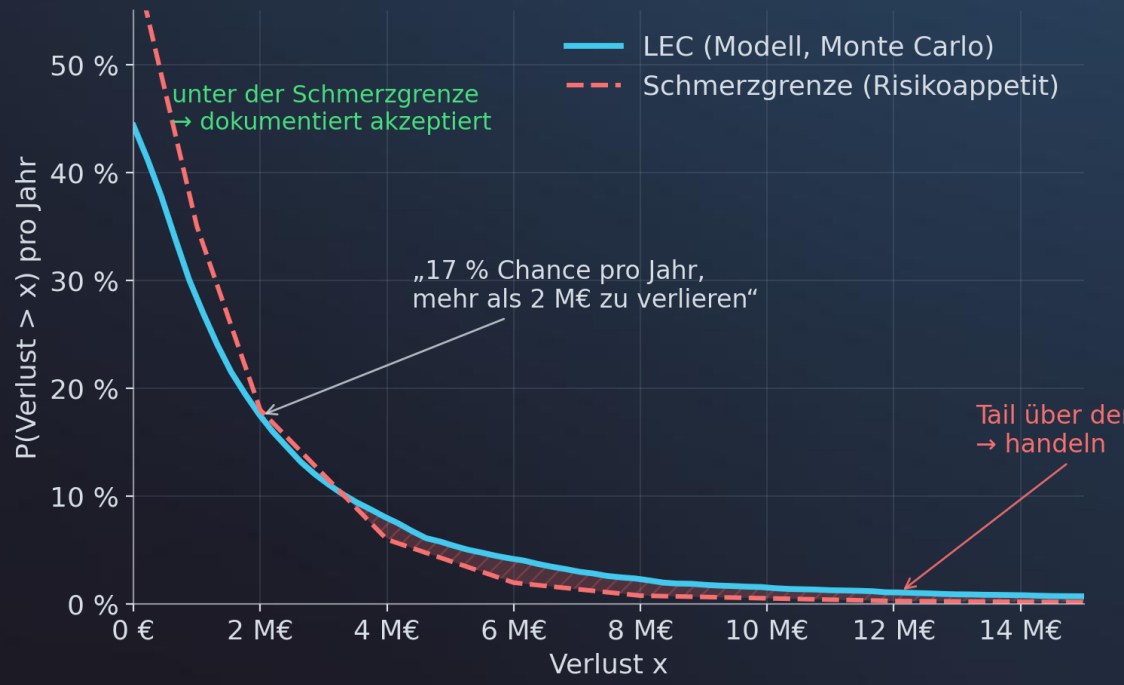


MC Output: Die Loss Exceedance Curve (LEC)

Jeder Punkt ist ein Satz: „5 % Chance pro Jahr, mehr als 10 M€ zu verlieren.“

Schmerzgrenze / Risikoappetit eventuell als zweite Kurve:

- Modell über der Toleranz → handeln
- darunter → dokumentiert akzeptiert



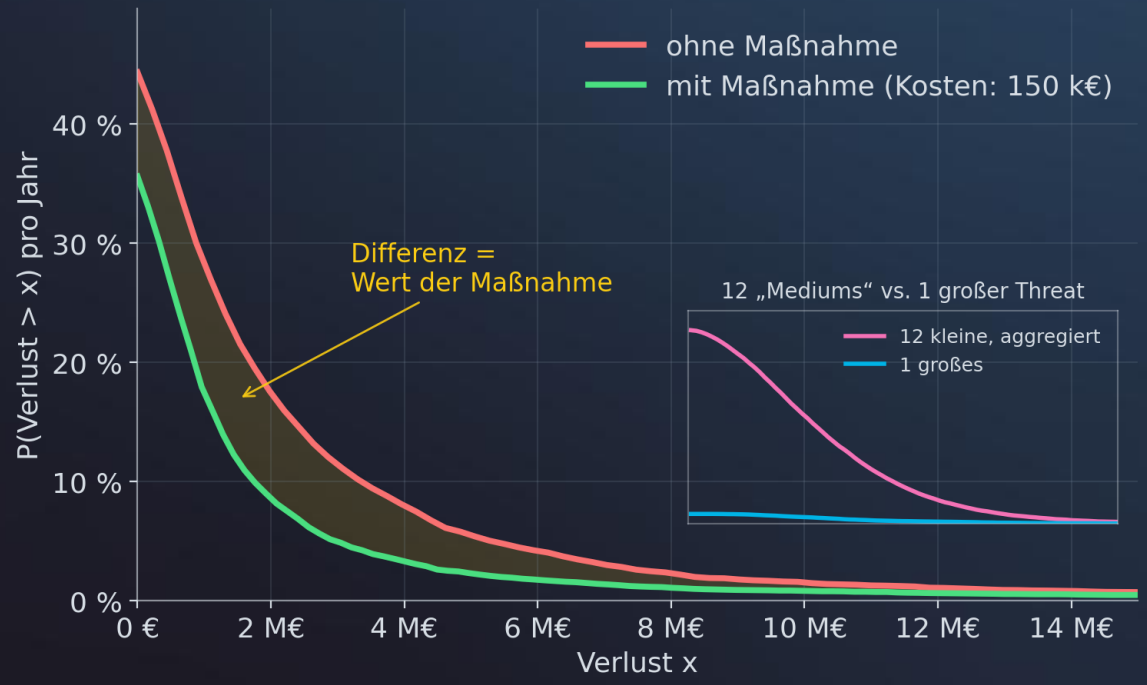
Was die LEC kann, was Matrix & Score nicht können

EIN Bild, alle Alternativen:

- LEC mit Maßnahme vs. ohne — Differenz sichtbar, Kosten daneben
- ROI-Vergleich ohne Statistikkenntnisse

Ehrliche Aggregation:

- Monte Carlo addiert VERLUSTE pro Trial — Geld + Geld ist erlaubt
- Matrix-Zellen und Scores kann man nicht addieren (Cox 2008)
- 12 „Mediums“ verschwinden im Raster …
- … aber ihr GEMEINSAMER Tail kann den einen „großen“ Threat übertreffen



Was Monte Carlo NICHT behauptet

Falsch: „Ihr werdet X verlieren.“

Richtig: Unser bestes gemeinsames Wissen als explizites Modell

Entscheidungsunterstützung:

- Alternativen vergleichbar machen
- Szenarien billig machen: „Was, wenn 10 % statt 5 %?“

Monte Carlo — Vorteile

Mit LECs: Ranking ✓, Vergleich ✓, Aggregation ✓

Die Modellparameter sind Abbildungen der Expertenschätzungen — eventueller Dissens wird SICHTBAR:

- Einigung → gemeinsamer Wert (oder Mittelwert)
- keine Einigung → beide Überzeugungen als Szenarien rechnen

Modelle müssen nicht perfekt sein, um Entscheidungen zu verbessern

Der Business Case fürs Threat Modeling

Das ewige Problem: kein Budget — für TM-Programm, Tooling, Maßnahmen

Der Weg:

1. Threats systematisch mit Expertenschätzungen erfassen
2. Einzelrisiken aggregieren → LEC des Programms / der Initiative
3. Vorher/Nachher-LEC + Kosten = finanzielles Argument in der Sprache des Managements

⇒ Aus „vertraut uns“ wird „hier ist die Kurve — und was es kostet, sie zu senken“.

Risquanter Register

Ein ABFRAGBARES, quantitatives Risikoregister

Experten definieren die Risikohierarchie mit ihren Schätzungen — ein Baum aus Portfolios und Leaves:

- Leaves = Risikoereignisse (Wahrscheinlichkeit + Verteilung)
- Portfolios aggregieren ihre Kinder (aggregierte Risikos)
- LEC auf JEDER Ebene des Baums

Risquanter Register

Inkrementell: nur geänderte Äste werden neu simuliert (Merkle-Cache) — Szenarien sind deshalb billig

„Abfragbar“ heißt: den LIVE-Baum befragen — kein SQL, sondern vage Quantoren („fast alle“, „rund die Hälfte“), wie ein CISO fragt

Open Source · work in progress: Kernfunktionalität production-quality, Ergonomie reift

github.com/risquanter/register

Was das Modell heute annimmt

Risquanter simuliert Risiken derzeit UNABHÄNGIG voneinander

Leitlinie: gemeinsame Ursachen als EIGENEN Knoten modellieren

- „Cloud Provider Outage“ = EIN Leaf mit vollem Blast Radius — nicht fünf heimlich gekoppelte kleine Leaves
- so denkt Threat Modeling ohnehin: gemeinsame Root Cause = ein Threat

Risikobasiert — regulatorisch sauber

Was die Regulierung verlangt — und was sie NICHT verlangt:

- CRA: Anforderungen gelten „based on the risks“; die Risikobewertung entscheidet, wie viel Sicherheit „genug“ ist (Fluchs, CRA Expert Group)
- NIS2 Art. 21: Maßnahmen nach „likelihood and severity of incidents“, all-hazards
- DORA: IKT-Risiko proportional zum Risiko

Nirgends steht: Geld oder Monte Carlo. Ein qualitativer Ansatz kann rechtlich reichen.

Risikobasiert — regulatorisch sauber

Aber: eine Threat-Liste mit L/M/H-Etiketten erbt genau die Schwächen von Matrix und Score — kein Maß, nicht aggregierbar.

Risquanter macht aus TM-Output echtes Risiko: pro Leaf Wahrscheinlichkeit + Schadensverteilung → aggregierte LEC

⇒ risikobasiert priorisieren im Sinne von CRA / NIS2 / DORA — sauber, vergleichbar, aggregierbar. Und mehr.

VQL: den Baum befragen

Vage Quantoren statt SQL — den LIVE-Baum befragen, wie ein CISO fragt (Fermüller et al. 2017)

Die CISO-Frage: „Tragen mindestens die HÄLFTE unserer Risiken in einem 1-von-20-Jahr einen Schaden über unserer Toleranz von $2M?“

Wort für Wort in eine Query:

- „mindestens die Hälfte“ → Q[>=]^{1/2}
- „unserer Risiken“ → leaf(x)
- „1-von-20-Schaden über $2M“ → gt_loss(p95(x), 2000000)

Ganze Query: Q[>=]^{1/2} x (leaf(x), gt_loss(p95(x), 2000000))

Antwort: erfüllt / nicht erfüllt — plus die Treffer. Keine Tabelle, kein SQL.

VQL live auf dem Demo-Baum

CISO-Frage 2 — Konzentration: „Wo sitzt unser Tail-Risiko — in der eigenen IT oder bei den Lieferanten?“

Erste Hälfte der Antwort — die eigene IT (Technology and Cyber):

- Q[>=]^{1/2} x (leaf_descendant_of(x, "Technology and Cyber"), gt_loss(p95(x), 2000000))

→ erfüllt (4/4): alle vier Cyber-Risiken tragen im 1-von-20-Jahr einen Schaden über $2M — die Schwelle war „mindestens die Hälfte“.

VQL live auf dem Demo-Baum

Gleiche Query, nur EIN Token wechselt (der Bereich) — jetzt die Lieferanten (Third-Party and Supply Chain):

- Q[>=]^{1/2} x (leaf_descendant_of(x, "Third-Party and Supply Chain"), gt_loss(p95(x), 2000000))

→ nicht erfüllt (0/3).

VQL live auf dem Demo-Baum

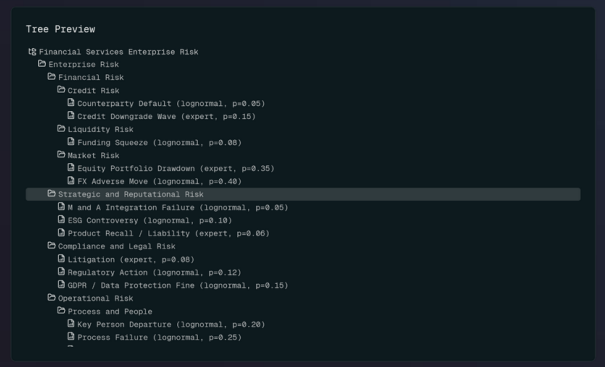
Das Tail sitzt in der eigenen IT, nicht bei den Lieferanten — in EINER Query belegt.

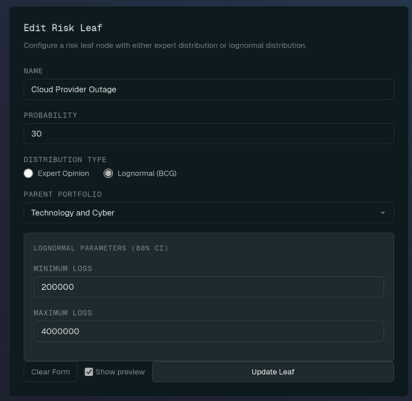
Ausblick — die Sprache kann mehr (dieselbe Frage, andere Form):

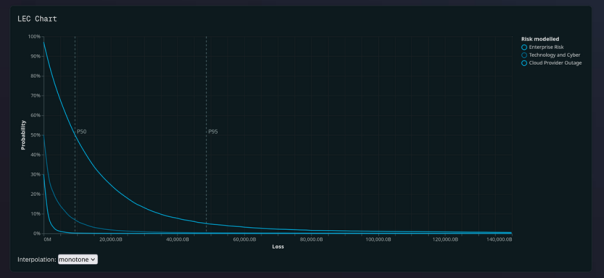
- Exceedance direkt: gt_prob(lec(x, 2000000), 0.05) — „>5 % Chance/Jahr über $2M?“
- echte vage Quantoren: Q[~]^{1/5} — „ungefähr ein Fünftel …?“
- Obergrenze statt Untergrenze: Q[<=]^{1/3} x (portfolio(x), …) — „höchstens ein Drittel der Portfolios über $50M?“ (Konzentration deckeln)
- verschachtelt (echte FOL): exists y . child_of(y, x) /\ … — „Portfolios mit mindestens EINEM Kind über $1M?“

THREAT MODELING CONNECT

Live-Demo







Selbst ausprobieren — Open Source & Ausblick

Jetzt: Register ist Open Source (AGPL v3) — github.com/risquanter/register

- Lokal in Minuten: docker compose --profile frontend up -d → http://localhost:18080

Bald: geplante öffentliche Demo-Instanz (ephemer, In-Memory)

Kommt:

- Szenario-Analyse: „Was-wäre-wenn“ als eigener Branch (O(1)), 3-Wege-Merge zurück ins Modell
- Time Travel: vollständige Historie als unveränderliche Commits — jeden früheren Stand ansehen oder zurückrollen
- weiter geplant: Abhängigkeitsgraph / „Company Configuration“

Wrap & Diskussion

Risiko ist **nicht** Priorität: eine sortierte Liste ist noch keine gute Investitionsentscheidung

Behauptet: vergleichbar · ehrlich aggregierbar · diskutierbar · budgetfähig

Nicht behauptet: die Zukunft vorherzusagen

Diskussion:

- „Ist das nicht alles Guesswork?“
- Äpfel vs. Birnen: Information Disclosure gegen DoS?
- Geld — oder Gain/Pain unterschiedlicher Akteure? (Ford Pinto)
- Wo würde das bei euch scheitern?

THREAT MODELING CONNECT

Referenzen

Budescu, D.V., Broomell, S.B., Por, H.-H. (2009). Improving Communication of Uncertainty in the Reports of the IPCC. Psychological Science 20(3), 299–308.

Cox, L.A. Jr. (2008). What's Wrong with Risk Matrices? Risk Analysis 28(2), 497–512.

Cox, L.A. Jr. (2009). What's Wrong with Hazard-Ranking Systems? An Expository Note. Risk Analysis 29(7), 940–948.

Cox, L.A. Jr. (2011). Clarifying Types of Uncertainty: When Are Models Accurate, and Uncertainties Small? Risk Analysis 31(10), 1530–1533.

Fermüller, C.G., Hofer, M., Ortiz, M. (2017). Querying with Vague Quantifiers Using Probabilistic Semantics. FQAS 2017, LNCS 10333.

FIRST.org (2019). CVSS v3.1: Specification Document. („CVSS Measures Severity, not Risk“)

Hubbard, D.W., Seiersen, R. (2016). How to Measure Anything in Cybersecurity Risk. Wiley.

Keelin, T.W. (2016). The Metalog Distributions. Decision Analysis 13(4), 243–277.

Spring, J., Hatleback, E., Householder, A., Manion, A., Shick, D. (2021). Time to Change the CVSS? IEEE Security & Privacy 19(2), 74–78.

Risquanter Register: github.com/risquanter/register

Verordnung (EU) 2024/2847 (Cyber Resilience Act), Art. 13 & Anhang I.

Richtlinie (EU) 2022/2555 (NIS2), Art. 21 — Risikomanagementmaßnahmen.

Verordnung (EU) 2022/2554 (DORA) — IKT-Risikomanagement.

Fluchs, S. FluchsFriction (Medium): „Cyber Resilience Act in 5 minutes“ u. a. — admeritia / EU CRA Expert Group.

Shostack, A. „Risk Management and Threat Modeling“, shostack.org.