OSMS ACADEMY

OSMS ACADEMY · GRUNDLAGEN

OSMS Fundamentals

Der offene Standard für Security-Kennzahlen - Begriffe, Prinzipien, Antworten.

Kostenloses Grundlagen-Material für CISOs, SOC-Teams, GRC & Audit.

opensecuritymetrics.org Stand Juli 2026 · frei teilbar

Security wird überall gemessen - nur nirgends gleich

Keine Vergleichbarkeit - Drei Teams, drei „Coverage“-Definitionen. Dieselbe Kennzahl bedeutet in jedem Haus - oft in jeder Abteilung - etwas anderes.

Zähe Audits - Jede Prüfung beginnt bei null: Was heißt die Zahl, woher kommt sie, warum diese Schwelle? Ohne normierte Definition ist jede Antwort Handarbeit.

Board ohne Belastbarkeit - Kennzahlen ohne Nenner, Mittelwerte ohne Verteilung, Grün ohne Datenvertrauen - Entscheidungen auf Schein-Präzision.

Die Ursache ist selten das Tooling - es fehlt die gemeinsame Definition. Genau dafür gibt es OSMS: einen offenen Standard, der festlegt, wie Security-Kennzahlen definiert, berechnet und berichtet werden.

Ein Standard, der Kennzahlen normiert - nicht noch ein Framework

Was - Offener, herstellerneutraler Standard für Security-Kennzahlen. Jede Kennzahl ist eine normative „Card“ mit Formel, Nenner, Schwellen-Guidance, Owner und Norm-Anker.

Warum - Schluss mit „jeder misst anders“: vergleichbar zwischen Teams und Firmen, auditierbar gegen Normen, board-tauglich durch Drill-down.

Wie - Maschinenlesbar (YAML/JSON), round-trip-validiert, Star-Schema furs Reporting. Der Standard ist Code, nicht nur Prosa.

DER 30-SEKUNDEN-PITCH

„OSMS definiert Sicherheitskennzahlen so, dass zwei Unternehmen unter derselben Zahl dasselbe verstehen - mit Formel, Nenner, Schwellen-Logik und Audit-Anker. Maschinenlesbar, framework-verankert, bis zur Rohdatenzeile erklärbar.“

Was OSMS NICHT ist: kein Tool · kein Framework-Ersatz (es operationalisiert NIST/ISO/CIS) · keine Compliance-Checkliste · kein Benchmark-Report

FÜR WEN?

Vier Perspektiven - ein gemeinsamer Standard

CISO & Management

Belastbare Zahlen fürs Board: ein Cockpit, dessen Werte bis zur Rohdatenzeile erklärbar sind - und Trends, die Definitionsänderungen überleben.

SOC- & Detection-Teams

Klare Messvorschriften statt Diskussionen: definierte Nenner, reproduzierbare Berechnungen, Guardrails gegen Fehlanreize.

GRC, Audit & Regulatorik

Jede Card trägt Norm-Anker (NIST CSF, ISO, CIS, DORA/NIS2) - Nachweise werden ableitbar statt jedes Mal neu verhandelt.

Berater & Hersteller

Ein offenes, maschinenlesbares Zielformat: Auswertungen, Dashboards und Produkte gegen einen Standard bauen statt gegen Hausdefinitionen.

FUNDAMENT

Die sechs Prinzipien - daran hängt alles

Nenner-Disziplin

Gemessen wird gegen eine autoritative Grundgesamtheit. Ohne vertrauenswürdigen Nenner ist jede Prozentzahl Fiktion.

Framework-Verankerung

Jede Card trägt ≥ 1 Norm-Anker (NIST CSF, ISO, CIS, Mitre ATT&CK, DORA/NIS2) - Audit-Verteidigung eingebaut.

Entscheidungs-Orientierung

Jede Card hat eine Decision Chain und eine management_question. Eine Kennzahl ohne Entscheidung ist Deko.

Ehrliche Statistik

P50/P90 statt Durchschnitt, direction statt „mehr ist besser“, Confidence-Gate statt Schein-Grün.

Anti-Gaming

Guardrails gegen Fehlanreize (Goodhart-Regel): Kennzahlen sollen steuern, nicht schmeicheln.

Maschinenlesbarkeit

YAML/JSON als Quelle der Wahrheit, Star-Schema, verlustfreie Reconciliation vom Board bis zur Rohdatenzeile.

BAUSTEINE

Card → Domäne → Katalog - und die Toolchain dahinter

Card

Die normative Einheit: EINE Kennzahl, vollständig definiert (47 Pflichtfelder). Stabile IDs mit Präfixen wie STD- / SOC- / AI- / HRM-.

Domain - Domäne

Themenfeld, das Cards bündelt - z. B. SOC – Detection & Response, Cloud Security, IAM, Governance & Compliance.

Catalog - Katalog

Die Gesamtheit aller Cards als maschinenlesbare Datei - über 300 Cards (Stand Juli 2026). Aktuelle Version: opensecuritymetrics.org.

DIE TOOLCHAIN

osms-catalog.yaml
Quelle der Wahrheit - alle Cards, alle Felder.

drill_spec.json
Drill-down-Mechanik je Archetyp (calculation_type).

star_schema.sql
Reporting-Schema mit GROUP-BY-Reconciliation.

drill_engine.py
Referenz-Implementierung - validiert den Round-Trip.

Merksatz: „Der Standard ist Code, nicht nur Prosa“ - wer OSMS implementiert, parst die YAML, statt ein PDF abzutippen.

Die sechs Kennzahl-Typen - wer misst was?

KPI

Key Performance Indicator → Leistungskennzahl

Misst Leistung & Ergebnis eines Prozesses. Frage: „Wie gut laufen wir?“

Bsp.: MTTR, Patch-SLA

KRI

Key Risk Indicator → Risiko-Frühindikator

Warnt, bevor es teuer wird - steigendes Risiko. Frage: „Wo brennt es bald?“

Bsp.: EOL-Exposure, stale Log Sources

KCI

Key Control Indicator → Kontrollindikator

Belegt Existenz & Wirksamkeit einer Kontrolle. Frage: „Greift die Kontrolle?“

Bsp.: Log Source Coverage

Outcome

Outcome Metric → Ergebniskennzahl

Das tatsächlich eingetretene Sicherheitsergebnis. Frage: „Was ist wirklich passiert?“

Bsp.: bestätigte Exfiltration

Confidence

Confidence Metric → Vertrauenskennzahl

Wie belastbar sind Daten & Nenner? Frage: „Trauen wir der Zahl?“

Bsp.: Inventar-Vollständigkeit

Maturity

Maturity Metric → Reifegradkennzahl

Wie erwachsen ist der Prozess (z. B. CMMI-Logik)?

Bsp.: PQC-Readiness

Lagging vs. Leading - beweisen vs. verhindern

NACHLAUFEND

Lagging

Misst, was passiert ist. Beweist Wirkung und Reaktionsfähigkeit - rückblickend.

Bsp.: MTTD, MTTR, Repeat-Incident-Rate

VORLAUFEND

Leading

Misst Voraussetzungen, bevor etwas passiert. Verhindert - vorausschauend steuerbar.

Bsp.: Coverage-KCIs, Baseline-Compliance, Scan-Kadenz

„Leading verhindert, Lagging beweist.“ Der Katalog ist bewusst ausbalanciert: Lagging-Kennzahlen belegen die Reaktion - die Leading-Schicht aus Coverage- und Denominator-KCIs macht Steuerung überhaupt erst möglich.

P0 / P1 / P2 und MVP-1/2/3 - wichtig vs. wann

P0

Priorität: MUSS

Kern-Set - ohne diese Cards kein glaubwürdiges Reporting.

P1

Priorität: SOLL

Wichtig - zweite Umsetzungsstufe nach dem Kern.

P2

Priorität: KANN

Ausbaustufe - Tiefe und Spezialfälle.

MVP - Minimum Viable Product → kleinster lauffähiger Stand. In OSMS ist MVP ein eigenes Card-Feld mit drei Ausbaustufen: MVP-1 = Go-live-Set, MVP-2 = zweite Welle, MVP-3 = Vollausbau. **Merksatz: P sagt, WIE WICHTIG eine Card ist - MVP sagt, WANN sie implementiert wird. Beide korrelieren stark, sind aber nicht identisch.**

Antwort auf „Womit starte ich?“: MVP-1-Set implementieren, P0 darin zuerst - live gehen, dann MVP-2/3 nachziehen.

P50 / P90 - warum der Durchschnitt lügt

Perzentil = der Wert, unter dem X % aller Fälle liegen. P50 = Median = „der typische Fall“. P90 = „der schlechte Tag“ - nur 10 % dauern länger.

Ø 6,6 h
Mittelwert - vom Ausreißer verzerrt

P50 = 4 h
Median - der typische Fall

P90 = 9 h
90 % sind schneller - SLA-Anker

Beispiel: 10 Incidents, Lösungszeit sortiert (h)

OSMS-Regel: Duration-Kennzahlen berichten P50 + P90 - nie nur den Mittelwert. SLAs gehören auf P90; der Ausreißer jenseits P90 bekommt einen eigenen Review. Gleiche Logik in der Risiko-Quantifizierung (FAIR): Verlustschätzungen als P50/P90-Spanne statt Punktwert.

Anatomie - die Pflichtbausteine jeder Card (echte Feldnamen)

1 Identität & Etiketten

id · name · domain · type · lag_type · priority · mvp · card_version · status · scope · unit

2 Sinn & Frage

purpose · definition · story · management_question („Welche Frage des Managements beantwortet die Card?“)

3 Rechnung

calculation · calculation_type (Archetyp) · calculation_plain · variables_inputs · numerator_denominator · rebuild_steps

4 Bewertung

target_thresholds · threshold_mode · direction · interpretation · special_cases_gates

5 Verantwortung & Takt

accountable_owner · escalation_owner · frequency

6 Daten & Vertrauen

data_sources · min_data_fields · data_confidence · confidence_production_rule · reproducibility

7 Wirkung & Anti-Gaming

decision_chain · next_chain · example_tasks · guardrails

8 Integrität & Audit

standards (Framework-Anker) · drilldown_lineage · version_break_rule

Definition of Done: Alle Bausteine gefüllt + Round-Trip-validiert = die Card ist „normativ“.

RECHNUNG

calculation_type - die Rechen-Archetypen

Ratio

Quote / Verhältnis

Zähler ÷ Nenner × 100. Der Arbeits-Archetyp - das Gros des Katalogs rechnet so.

Bsp.: Coverage %

Composite

zusammengesetzter Index

Mehrere Sub-Scores zu einem Wert verdichtet.

Bsp.: EDR Effectiveness

Duration

Zeitdauer

Zeitspannen - berichtet als P50/P90, nie nur Ø.

Bsp.: MTTR, Dwell Time

Weighted Sum

gewichtete Summe

Teilwerte × Gewichte, aufsummiert.

Bsp.: gewichteter Zielwert

Count

Anzahl

Absolute Zählung - oft mit zero_is_target.

Bsp.: offene P0-Findings

Weighted Average

gewichteter Durchschnitt

Mittelwert mit Gewichtung je Element.

Bsp.: gewichtete Reife

Zusammen 13 calculation_types: die sechs oben plus Unit Cost, Delta, Penalty, Index, Score, Monetary Risk, Ranking. Das Drill-Framework mappt alle auf 6 generische Mechaniken (ratio, duration, count, delta, component_tree, ranking) - der Archetyp bestimmt die Drill-down-Logik in drill_engine.py.

PRINZIP NR. 1

Numerator ÷ Denominator - der Nenner ist der Standard

Numerator = Zähler (das Gemessene) · **Denominator = Nenner** (die autoritative Grundgesamtheit, gegen die gemessen wird - „authoritative population“)

Effective Log Coverage (%) = onboarded **UND** healthy ÷ In-Scope-Assets laut Logging-Baseline × 100

Woher kommt der Nenner? - Inventar (CMDB) × Policy/Baseline: Welche Asset-Klasse MUSS was liefern? Das ist das autoritative „Soll“.

Zerlegung statt Konkurrenz - Ebenen komponieren auf die Headline: Depth × Health = 50 % × 97,5 % = 48,75 % - exakt 3.900 ÷ 8.000.

Wer bewacht den Nenner? - Confidence-Metriken (z. B. Inventar-Vollständigkeit) - sie messen, ob der Grundgesamtheit zu trauen ist.

„Ohne vertrauenswürdigen Nenner ist jede Prozentzahl Fiktion.“

direction & threshold_mode - was heißt „gut“?

higher_is_better	lower_is_better	context_dependent	zero_is_target	band_is_better
höher = besser	**niedriger = besser**	**nur mit Kontext**	**Ziel ist 0**	**Zielband**
Coverage, Abdeckung	*MTTR, Backlog*	*Volumen, Kosten*	*überfällige P0-Findings*	*Über- & Untersteuerung kritisch*

threshold_mode - das Schwellen-Prinzip: OSMS schreibt KEINE universellen Schwellenwerte vor. Der Leitfaden liefert Guidance - verbindlich werden Schwellen erst durch lokale Bindung an Risikoappetit, Servicekritikalität und Datenvertrauen. Helper-Cards sind nicht eigenständig berichtsfähig; sie erben Schwellen von der Haupt-Card.

Antwort auf „Welchen Zielwert schreibt OSMS vor?“: Keinen festen - Guidance ja, Bindung an den eigenen Risikoappetit ist Pflicht. Ein Standard, der überall 95 % verlangt, wäre unseriös.

RAG-Logik & das Data-Confidence-Gate

RAG = Red / Amber / Green - Ampellogik je Card (target_thresholds). Beispiel-Guidance:

Scope	Grün	Amber	Rot
Kritische Assets	≥ 95 %	85–94 %	< 85 %
Gesamtbestand	≥ 90 %	80–89 %	< 80 %

Wichtig: Das sind Beispiel-Bänder (Guidance) - lokal an Risikoappetit binden (threshold_mode).

Plus direction: Bei lower_is_better dreht sich die Ampel - die Logik bleibt gleich.

DATA-CONFIDENCE-GATE

Grün gibt es nur mit belastbaren Daten.
Felder: data_confidence + confidence_production_rule + min_data_fields. Die Card wird nur „produziert“, wenn Mindest-Datenfelder und Confidence-Kriterien erfüllt sind - sonst gilt sie als nicht belastbar (kein Schein-Grün).

Typische Frage: *„Der Wert ist doch gut - warum ist die Ampel nicht grün?“*
Antwort: Weil der Nenner wackelt. Erst Datenvertrauen herstellen, dann Farbe.

Decision Chain, next_chain & Guardrails

DECISION CHAIN - ENTSCHEIDUNGSKETTE

Grün → beobachten
Trend & Kontext im Blick behalten - keine Aktion nötig.

Amber → Owner + Frist
accountable_owner handelt, escalation_owner ist informiert.

Rot → Eskalation
Ticket, benannter Owner, Tracking bis Closure - vordefiniert, nicht improvisiert.

next_chain → Folgefrage
Welche Kennzahl/Frage kommt danach? Kennzahlen verketten sich zur Steuerung.

„Eine Kennzahl ohne Entscheidung ist Deko."

GUARDRAILS - LEITPLANKEN GEGEN GAMING

Onboarding nie ohne Health berichten - sonst zählen „connected but silent"-Quellen als Erfolg.

Utilization nur gegen kuratierten, security-relevanten Satz - sonst bestraft die Quote gesprächige Quellen.

„Genutzt" heißt: in einer aktiven, getesteten Regel referenziert - sonst lädt die Quote zu Dummy-Regeln ein.

Goodhart-Regel: „Wird die Kennzahl zum Ziel, hört sie auf, gut zu messen." Guardrails sind die eingebaute Gegenmaßnahme.

AGGREGATION

Vom Event zum Board - und verlustfrei zurück

1

Rohdaten / Events

SIEM, EDR, Cloud-Logs - die unterste Wahrheit.

2

Operative Cards

Einzelkennzahlen je Domäne (SOC-/STD-/AIM-/HRM-IDs).

3

Domänen-Rollup

Verdichtung je Themenfeld - gleiche Formel-Logik.

4

Board-Cockpit (L4)

9 kanonische KPIs - die Management-Sicht.

Reconciliation-Garantie: Jede aggregierte Zahl zerfällt verlustfrei in ihre Bestandteile (GROUP-BY-Garantie im Star-Schema). drilldown_lineage dokumentiert den Pfad je Card - jede Board-Zahl ist bis zur Rohdatenzeile erklärbar. Das ist der Unterschied zwischen einem Dashboard und einem Standard.

Board-Frage „Woher kommt diese Zahl?“ → Antwort in Sekunden statt Sonderanalyse.

VERANKERUNG

standards - die Framework-Anker jeder Card

NIST CSF 2.0

Funktionsmodell GV · ID · PR · DE · RS · RC - das Rückgrat des Mappings.

ISO/IEC 27001

ISMS-Anker - Managementsystem & Controls (Annex A).

CIS Controls v8.1

Konkrete, priorisierte Kontrollen - die operative Brücke.

ISO/IEC 27004

Die Mess-Methodik-Norm - wie man Sicherheit misst.

NIS2 / DORA

EU-Pflichten: Meldefristen (24h/72h), Register, TLPT, Resilienz.

MITRE ATT&CK

Techniken & Data Sources - Basis für Detection-Coverage.

Funktion: Jede Card trägt ≥ 1 Anker im Feld standards → eingebaute Audit-Verteidigung. OSMS ersetzt diese Rahmenwerke nicht - es macht messbar, ob ihre Anforderungen erfüllt sind. In Arbeit: CRA, EU AI Act.

STAND & WEG

Ein lebender Standard - Version, Wellen, Stabilität

v0.9

Basis-Release

Kern-Katalog, Domänenstruktur, Toolchain — Public Review bis 31. Okt 2026.

v1.0

Fundament + Regulatorik

Coverage-/Denominator-Schicht, DORA- & NIS2-Pflichtkennzahlen.

Ziel: Ende November 2026

v1.1

Operative Tiefe

Detection-, Daten-, Netzwerk-, Lieferketten- und Resilienz-Familien.

v1.2

Scope-Erweiterung

AppSec/CRA, OT/ICS, Privacy-Ops - als bewusste Scope-Entscheidungen.

Stabilitätsversprechen: Card-IDs werden nie umnummeriert; Definitionsänderungen brechen Trends kontrolliert (version_break_rule) - Investitionen in OSMS bleiben gültig.

Konformitäts-Annex: Meta-Metriken prüfen das Mess-System selbst (Nenner-Drift, Confidence-Gate, Owner-Attestierung) - OSMS ist ein selbst-prüfender Standard.

Der Katalog wächst: über 300 Cards (Stand Juli 2026). Aktuelle Version, Changelog und Roadmap: opensecuritymetrics.org.

LOSLEGEN

In vier Schritten zum ersten OSMS-Reporting

1 · Katalog laden

kpi_catalog.yaml - maschinenlesbar, versioniert, mit allen Pflichtfeldern.

2 · Startpaket wählen

MVP-1-Set, P0 zuerst - das Go-live-Set fürs erste Reporting.

3 · Nenner & Schwellen festziehen

Inventar × Policy als Nenner; Schwellen an den Risikoappetit binden.

4 · Berichten & steuern

Fester Takt - mit Confidence-Gate, Decision Chain und Drill-down.

Offen & kostenlos. OSMS ist herstellerneutral, maschinenlesbar und frei zugänglich. Katalog, Dokumentation, Nutzungsbedingungen und Updates: **opensecuritymetrics.org**

Fragen, Feedback oder Interesse an der OSMS Academy? → Kontakt über die Webseite.

WENN SIE NUR FÜNF SÄTZE MITNEHMEN

1. „Der Nenner ist der Standard.“
2. „Leading verhindert, Lagging beweist.“
3. „P50 ist der Alltag, P90 der schlechte Tag.“
4. „Eine Kennzahl ohne Entscheidung ist Deko.“
5. „Grün gibt es nur mit belastbaren Daten.“

Katalog, Doku & Updates: **opensecuritymetrics.org**