SQUR

FACHBEITRAG ENERGIEWIRTSCHAFT

Autonome Pentests für Energieversorger: Sicherheit im Takt Ihrer Releases

Wie Stadtwerke und Versorger ihre Portale und APIs im risikogerechten Takt und nach wesentlichen Änderungen prüfen lassen, Schwachstellen schneller schließen und das gegenüber Geschäftsleitung und Aufsicht belegen.

Adam Lundqvist, Gründer und CEO von SQUR · LinkedIn · squr.ai Rechtsstand und Quellenabruf: 8. November 2026

Das Wichtigste in Kürze

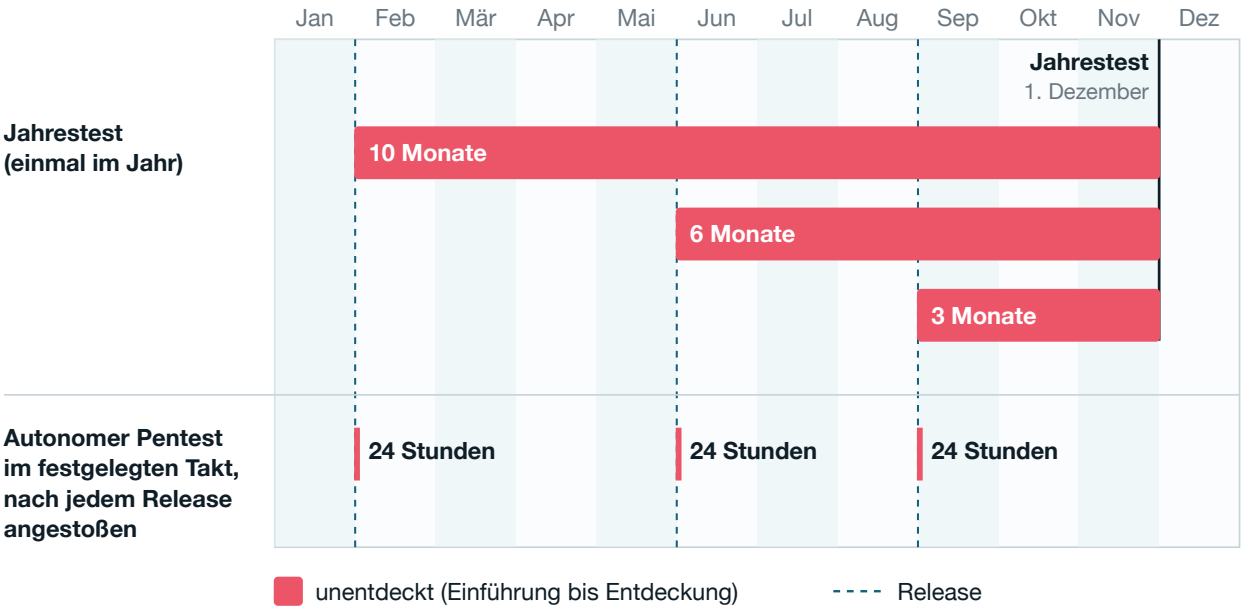
So kann es aussehen. Ihr Kundenportal bekommt ein Release, und am nächsten Tag liegt ein Pentest-Ergebnis zu genau diesem Stand vor. Wer Sicherheit verantwortet, erhält reproduzierbare Befunde mit Beweis und kann sie direkt zuweisen. Die Entwicklung prüft per Retest, ob die Korrektur die Lücke wirklich schließt. Die Geschäftsleitung sieht auf einen Blick, welche wichtigen Anwendungen einen aktuellen Prüfnachweis haben. Und Sie können weitere Anwendungen regelmäßig prüfen lassen, ohne jeden Test selbst durchführen zu müssen; Ihr Team hält Testkonten und Testdaten aktuell und verantwortet die Behebung.

- **Die entscheidende Größe ist das Expositionsfenster:** die Zeit, die eine ausnutzbare Schwachstelle unentdeckt offen ist. Beim jährlichen Pentest sind das im Mittel sechs Monate, bei einem Test nach jedem Release wenige Tage plus Ihre Behebungszeit, die damit zum Engpass wird.
- **Gefährlich sind in Portalen und APIs von Versorgern vor allem Berechtigungsfehler:** fremde Zählerstände, übernommene Kundenkonten, erschlichene Vorteile nach § 14a EnWG. Sie berühren direkt Geld, Kundendaten und Vertrauen. Finden lassen sie sich mit Tests, die mehrere Identitäten gegeneinander führen.
- **Autonome Pentests machen diesen Takt über ein ganzes Portfolio praktikabel.** Schon bei 100 Anwendungen und APIs ergeben sich über 200 Testläufe im Jahr. Testidentitäten und Regeln richten Sie einmal ein und halten sie aktuell; die Tests selbst laufen ohne Ihr Zutun.
- **Empfehlenswert sind sie, wenn ihre Grenzen außerhalb des Agenten gelten.** Mit drei einfachen Abnahmetests prüfen Sie das selbst.
- **Das Gesetz verlangt dokumentierte Wirksamkeit** (§ 30 BSIG, § 5d EnWG) und verpflichtet die Geschäftsleitung zur Überwachung. Pentest-Berichte mit Retest sind dafür ein konkreter technischer Beleg.
- **Fangen Sie mit einer Anwendung an.** Der erste Lauf bringt das Programm in Gang; erweitern lässt es sich jederzeit, verlorene Zeit nicht.

Einleitung: Sicherheit im Takt Ihrer Releases

Ein Pentest beantwortet die Frage, was an einem bestimmten Tag in einem bestimmten Umfang ausnutzbar war. Das bleibt wertvoll. Kundenportale und Apps, Backends für Ladepunkte, Schnittstellen für Einspeiser und Direktvermarkter und Portale für steuerbare Verbrauchseinrichtungen ändern sich aber im Takt der Softwareentwicklung, häufig wöchentlich. Zwischen zwei jährlichen Prüfterminen entstehen so Zeiträume, in denen niemand geprüft hat, was sich geändert hat.

Abbildung 1 zeigt das an einem Beispiel mit drei Releases im Jahr und einem jährlichen Pentest im Dezember: Die Releases erzeugen unentdeckte Zeiträume von zehn, sechs und drei Monaten. Prüft ein Autonomer Pentest nach jedem größeren Release und liegt das Ergebnis nach 24 Stunden vor, schrumpft dieser Zeitraum für jede Schwachstelle, die der Lauf findet, auf diese 24 Stunden. Die Stellschraube ist der Abstand zwischen Änderung und Prüfung.



Maßstäblich: 24 Stunden sind rund 1/30 Monat und im Raster kaum sichtbar. Schematisch, Beispieltermine. Die Zeit bis zur Behebung kommt in beiden Fällen hinzu.

Abbildung 1: Expositionsfenster, schematisch. Beispielhafte Releases im Februar, Juni und September, Jahrestest im Dezember; darunter ein Autonomer Pentest im festgelegten Takt, der binnen 24 Stunden abgeschlossen ist. Annahmen: Die betrachtete Schwachstelle wird im nächsten Lauf gefunden, beide Verfahren decken sie gleichermaßen ab; die Behebungszeit kommt in beiden Fällen hinzu.

Gegenstand dieses Beitrags sind Webanwendungen und APIs, auch dort, wo sie an Netzleittechnik und OT angebunden sind. Er richtet sich an die Person, die im Versorger Sicherheit verantwortet, und liefert ihr auch die Argumente für die Geschäftsleitung. Zum Mitnehmen gibt es eine Einseiter-Checkliste (Anhang A) und ein Rechenblatt für das Expositionsfenster (Anhang B).

Ein Lauf von Anfang bis Ende

Wie das im Alltag aussieht, zeigt ein Beispiel. Es beschreibt den Ablauf, wie er vorgesehen ist; Anwendung, Rollen und Befund sind erfunden, nicht gemessen.

Ein Stadtwerk bringt am Dienstag eine neue Version seines Kundenportals in Betrieb; derselbe Stand läuft in der Testumgebung. Neu ist, dass Kunden ihren Abschlag selbst anpassen können. Am Mittwoch liegt das Ergebnis des Autonomen Pentests vor:

Schritt	Was passiert
Umfang	Portal und API in der Testumgebung, Freigabeliste mit Hosts und Pfaden. Schreibende Tests nur auf Testverträgen, zerstörerische Handlungen ausgeschlossen.
Identitäten	Zwei Testkunden (A und B) und ein Mitarbeiterkonto, einmal eingerichtet und seitdem für jeden Lauf genutzt.
Erkundung	Der Agent findet die neue Funktion "Abschlag ändern", erkennt, dass sie eine Vertragskennung übergibt, und probiert sie mit den anderen Testidentitäten.
Befund	Kunde B kann den Abschlag von Kunde A ändern. Die Gegenprobe als A zeigt den geänderten Betrag.
Bedeutung	Ein Angreifer kann Abschläge beliebiger Kunden setzen, wenn er deren Vertragskennung kennt; sie steht auf jeder Rechnung.
Beweis	Anfrage von B auf den Vertrag von A, Antwort des Portals, Zustand vorher und nachher, geschwärzt und reproduzierbar.
Ihr Anteil	Testkonten und Freigabeliste einrichten und aktuell halten; Befunde sichten und zuweisen; Korrektur und Retest anstoßen.
Behebung und Retest	Die Entwicklung ergänzt die Prüfung, ob der Vertrag dem Aufrufer gehört. Der Retest am Donnerstag wiederholt genau die belegte Anfrage und bestätigt: geschlossen. Noch am Donnerstag geht die Korrektur in Betrieb.
Grenzen	Nicht geprüft: Zahlungsdienstleister außerhalb der Freigabeliste, Produktivdaten, Lastverhalten. Der Bericht nennt das ausdrücklich.

Im Jahresrhythmus würde eine im nächsten Test erkennbare Lücke im Mittel erst nach sechs Monaten entdeckt. In diesem Beispiel liegt der Befund am nächsten Tag vor, einen Tag später besteht die Korrektur den Retest und ist in Betrieb. Den Testfall "ein Kunde ändert nur seinen eigenen Vertrag" musste niemand vorgeben: Der Agent erkundet Funktionen und Objektbeziehungen und entwickelt daraus Prüfhypothesen. Fachliche Sonderregeln ergänzen Sie, etwa ob ein Installateur Verbrauchsdaten sehen darf oder ob Haushaltsmitglieder auf einen gemeinsamen Vertrag zugreifen dürfen. Der Bericht zeigt, welche Funktionen und Identitäten tatsächlich geprüft wurden und welche Bereiche nicht erreicht wurden.

Wie ein echter Lauf mit Befund, Beweis und Retest aussieht, zeigen wir Ihnen im Gespräch an einer Testanwendung, auf Wunsch anschließend an einer Ihrer Anwendungen: **squr.ai/meeting**

Warum jetzt: Was öffentliche Schwachstellendaten zeigen

Der direkteste Grund für häufigere Prüfung ist Ihr eigener Release-Takt: Was sich ändert, sollte geprüft werden. Öffentliche Schwachstellendaten zeigen zusätzlich, dass das Umfeld schneller wird. Grundlage sind die neu veröffentlichten CVEs je Monat in der National Vulnerability Database der USA (NVD, ohne zurückgezogene Einträge) und alle ICS-Advisories im CSAF-Archiv der US-Behörde CISA, über die Hersteller von Netztechnik, Wechselrichtern und Ladeinfrastruktur ihre Lücken melden, jeweils bis Ende Oktober 2026.

Die Menge wächst. 2025 wurden 48.150 CVEs veröffentlicht, von Januar bis Oktober 2026 bereits 89.846. Von 2021 bis 2025 verdoppelte sich die monatliche Zahl etwa alle drei Jahre, seit April 2026 etwa alle 3,7 Monate (exponentieller Fit April bis Oktober 2026, Bestimmtheitsmaß 0,99). Im Oktober 2026 waren es 17.052, das Vierfache des Monatsmittels von 2025. Abbildung 2 schreibt beides fort: das Szenario, dass der Oktoberwert konstant bleibt, und das Szenario, dass die Verdopplung anhält. Beides sind Annahmen, keine Prognosen.

Der Anteil der Industrie-Meldungen sinkt. Die ICS-Advisories der CISA enthielten 2024 2.008 neue CVEs, 2025 1.789 und von Januar bis Oktober 2026 1.927. Ihr Anteil an allen neuen CVEs fiel von 5,0 Prozent (2024) auf 2,1 Prozent (2026 bis Oktober), vor allem weil die Gesamtzahl wächst. Von den CVEs, die CISA dem Sektor Energie zuordnet, sind 60 Prozent laut CVSS über das Netz ausnutzbar; nur rund 6 Prozent setzen Nahzugriff oder physischen Zugang voraus.

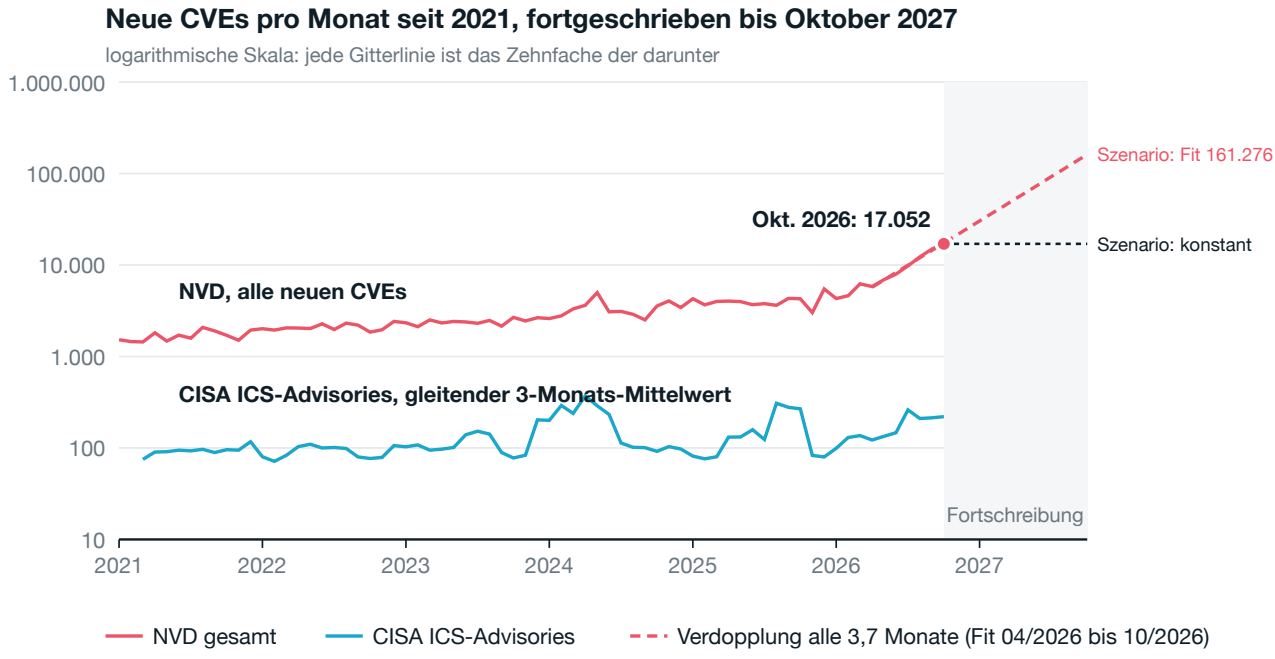


Abbildung 2: Neue CVEs pro Monat in der NVD und in den ICS-Advisories der CISA (CVE im ersten Advisory gezählt, gleitender 3-Monats-Mittelwert), Januar 2021 bis Oktober 2026, logarithmische Skala. Gestrichelt: zwei Szenarien bis Oktober 2027, Fortschreibung des Fits April bis Oktober 2026 und Oktoberwert konstant; beides Annahmen, keine Messung. Quelle: NVD API 2.0, cisagov/CSAF, eigene Auswertung.

Was das für Ihr Programm heißt. Die weltweite Zahl sagt nicht, wie viele dieser Schwachstellen Ihre Systeme betreffen. Sie zeigt aber, dass sich das Umfeld schneller bewegt als ein jährlicher Prüfrhythmus. Und das Wichtigste steht in keinem dieser Feeds: Eine Berechtigungslücke in Ihrem eigenen Kundenportal bekommt nie eine CVE-Nummer. Sie finden sie nur durch eigene Prüfung, und deren Takt bestimmen Sie selbst (Kapitel 2).

Grenzen der Daten. Die NVD zählt veröffentlichte, nicht ausgenutzte Schwachstellen, und ein Teil des Anstiegs ist Erfassung statt neuer Schwäche. Die ICS-Auswahl ist die der CISA; ein Netzvektor besagt, dass ein Netzzugang genügt, nicht, dass das Gerät im Internet steht.

Drei Verfahren und wofür sie stehen

	Schwachstellenscanner	Pentest durch Menschen	Autonomer Pentest
Stärke	Breite: bekannte Schwachstellen, Konfiguration, Komponentenstände	Erfahrung und Kontext; als Red Teaming Übungen über die ganze Organisation	Tiefe und Takt: Angriffsketten über mehrere Identitäten, wiederholbar über das ganze Portfolio
Wie	Signaturen und Versionsabgleich	Erkunden, Hypothesen, gezieltes Ausprobieren	Ein Agent nutzt Scanner als einen Schritt, erkundet, bildet Hypothesen und führt mit mehreren Identitäten Angriffsschritte aus
Berechtigungs- und Logikfehler	Selten, ohne passende Identitäten und Regeln	Ja, im Rahmen des Zeitbudgets	Ja, im festgelegten Umfang
Nachweis	Trefferliste, oft mit Fehlalarmen	Pentest-Bericht mit Beweisen	Pentest-Bericht mit Exploit-Beleg je Schwachstelle und Retest
Takt	Laufend	Planung mit Wochen Vorlauf, meist jährlich	Nach jedem Release oder im festgelegten Takt, Ergebnis in 24 Stunden
Aufwand für Sie	Gering	Scoping, Testkonten und Begleitung je Termin	Einmalig Freigabeliste, Testidentitäten und Abbruchweg; danach Ergebnisse sichten und Konten aktuell halten

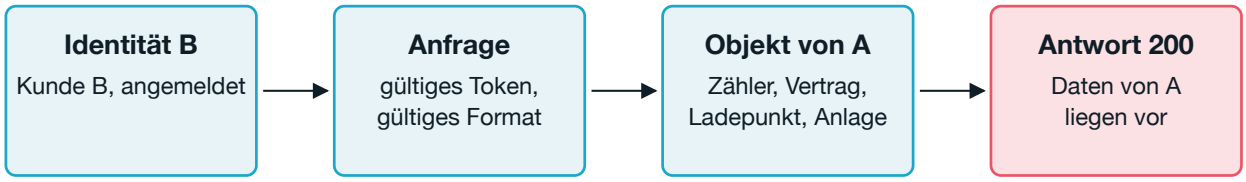
Die Verfahren bauen aufeinander auf. Ein Autonomer Pentest setzt Scanner selbst als einen Arbeitsschritt ein und geht dann weiter: Er verknüpft Funde zu Angriffsketten, prüft Berechtigungen mit mehreren Identitäten und belegt jede Schwachstelle mit einem Exploit. Fest programmierte Berechtigungstests, wie sie OWASP für jedes Release empfiehlt, sind eine gute Ergänzung, prüfen aber nur, was vorher jemand aufgeschrieben hat; der Autonome Pentest erkundet auch, was niemand vorhergesehen hat. Red Teaming hat seinen Platz bei Gesamtübungen, die Angriff und Verteidigung über die ganze Organisation prüfen, mit Erkennung, Reaktion, Menschen und Prozessen, statt einzelne Anwendungen.

Kapitel 1: Wo es bei Versorgern ernst wird, fünf Angriffsketten

1.1 Das gemeinsame Muster

Die fünf Ketten betreffen Zählerstände, die Verknüpfung von Verträgen, Ladepunkt-Backends, Einspeise-APIs und steuerbare Verbrauchseinrichtungen. Es sind Beispiele für typische Fehlerbilder, keine dokumentierten Vorfälle. Ihre Struktur ist dieselbe: Eine angemeldete Anfrage ist technisch gültig, aber fachlich nicht erlaubt. Sitzung, Token und Format stimmen, die Antwort ist ein Statuscode 200. Falsch ist allein, dass das Objekt dem Aufrufer nicht gehört, dass die Rolle die Funktion nicht aufrufen dürfte oder dass der Zustandsübergang nicht vorgesehen war. Die Referenz ist die OWASP API Security Top 10 (2023), vor allem API1 bis API3 und API5.

Ein Scanner ohne passende Identitäten sieht darin eine normale Antwort. Finden lassen sich solche Fehler mit einem Test, der mehrere Identitäten gegeneinander führt und die Antwort fachlich bewertet. Das Grundmuster lautet: Darf diese Identität dieses Objekt mit dieser Funktion berühren?



Technisch gültig, fachlich nicht erlaubt: Sitzung, Token, Format und Statuscode stimmen.

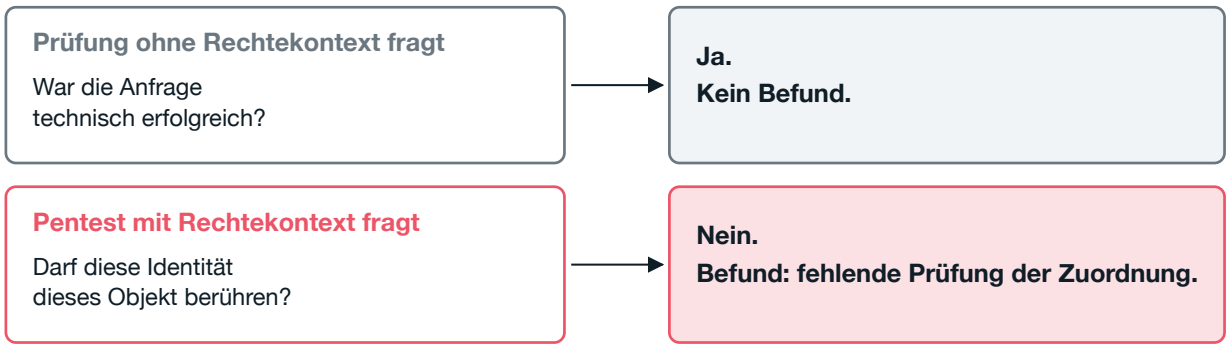


Abbildung 3: Gemeinsame Struktur der fünf Angriffsketten. Die Anfrage ist technisch gültig; erst die Frage nach der Zuordnung von Objekt und Aufrufer zeigt den Fehler.

Dafür braucht der Test von Ihnen **zwei bis drei Testidentitäten je Rolle** (Kunde A, Kunde B, Installateur, Mitarbeiter) mit eigenen Testverträgen, Testzählern oder Testanlagen, auf die sich schreibende Tests beschränken. Die **Matrix Rolle x Funktion x Objekt** (Abbildung 4) müssen Sie nicht vorgeben: Ein Autonomer Pentest erkundet Funktionen und Objektbeziehungen, entwickelt daraus Prüfhypothesen und führt die gefundenen Funktionen mit fremden Identitäten erneut aus. Fachliche Sonderregeln ergänzen Sie, etwa gemeinsame Verträge oder delegierte Zugriffe, denn das beobachtete Verhalten allein zeigt nicht, was erlaubt sein soll. Der Bericht weist aus, welche Kombinationen geprüft wurden. Dabei gilt: Eine Kennung, die auf der Rechnung steht, ist als öffentlich zu behandeln. Belegt ist ein Befund erst als Paar: die Anfrage von B auf das Objekt von A, plus Gegenprobe mit der richtigen Rolle.

Objekt: Zähler und Abschlagsplan von Kunde A. Alle Zellen sind ein Beispiel.

	Stände lesen	Stand melden	Abschlag ändern	Plan widerrufen	Vertragsdaten lesen
Kunde A (Eigentümer)	erlaubt, wie erwartet	erlaubt, wie erwartet	erlaubt, wie erwartet	erlaubt, wie erwartet	erlaubt, wie erwartet
Kunde B	verweigert	verweigert	**erlaubt, aber fachlich unzulässig**	verweigert	**erlaubt, aber fachlich unzulässig**
Installateur	verweigert	**erlaubt, aber fachlich unzulässig**	verweigert	verweigert	verweigert
Mitarbeiter	erlaubt, wie erwartet	erlaubt, wie erwartet	erlaubt, wie erwartet	verweigert	erlaubt, wie erwartet

erlaubt, Ergebnis wie erwartet erlaubt, aber fachlich unzulässig: Befund verweigert

Jede Funktion wird mit jeder fremden Identität erneut ausgeführt; bewertet wird die Zuordnung, nicht der Statuscode.

Abbildung 4: Testmatrix Rolle x Funktion x Objekt. Alle Zellen sind ein Beispiel, keine Messung.

1.2 Kette 1: Zählerstände und Abschläge fremder Kunden

Ein Portal oder eine App hat vorhersagbare oder aus Rechnungen ableitbare Objektkennungen, und die Prüfung, ob das Objekt dem Aufrufer gehört, fehlt oder greift nur im Frontend. Kunde B liest mit der Zählerkennung von A Verbrauch und Vertragsdaten, meldet einen Zählerstand für den Zähler von A oder ändert dessen Abschlag. Wird das ausgenutzt, kann daraus eine meldepflichtige Verletzung des Schutzes personenbezogener Daten werden (Art. 33 DSGVO), und Abrechnungen beruhen auf manipulierten Werten, sofern keine nachgelagerte Plausibilisierung greift. Der Test liest und schreibt mit fremder Kennung, prüft danach als A, ob sich etwas verändert hat, und sendet Felder, die das Frontend nie sendet (Abschlagsbetrag, Status, Tarif).

1.3 Kette 2: Kontoübernahme über die Vertragsverknüpfung

Die Funktion "Vertrag verknüpfen" verlangt als einzigen Besitznachweis Kundennummer und Zählernummer, beide stehen auf der Rechnung. Unterscheiden die Fehlermeldungen gültige von ungültigen Kombinationen, wird der Endpunkt zum Orakel. Lässt sich nach der Verknüpfung die Kontakt-E-Mail ändern und ein Passwort-Reset auslösen, ist die Übernahme vollständig, mit Zugriff auf IBAN, Umzug, Kündigung und Wechsel. Ein Angreifer könnte dann Lieferverträge kündigen oder Gutschriften auf sein eigenes Konto umleiten, und der Versorger erführe davon oft erst durch die Beschwerde des Kunden. Der Test prüft das Orakelverhalten an eigenen Testverträgen, nie durch Iteration gegen echte Kunden, und ob für kritische Änderungen eine zweite Bestätigung erzwungen wird.

1.4 Kette 3: Ladepunkt-Backends

Das Backend nimmt WebSocket-Verbindungen von Ladepunkten an, meist nach OCPP, oft mit der Ladepunkt-Kennung in der Verbindungsadresse. Zu prüfen ist, welches Sicherheitsprofil das Backend tatsächlich durchsetzt, ob eine zweite Verbindung mit derselben Kennung die erste verdrängt, ob sich

ein unbekannter Ladepunkt registrieren lässt und ob Nutzer B eine Ladung von Nutzer A einsehen oder stoppen kann. Im Ernstfall fallen Ladepunkte aus oder Ladevorgänge landen auf der Rechnung fremder Nutzer, beides trifft Umsatz und Ruf des Betreibers unmittelbar. Das geschieht ausschließlich in einem Testmandanten mit Testladepunkten oder Simulator, nie im Echtbetrieb. Ob ein Anbieter Ladepunkt-Protokolle testen kann, lassen Sie sich an einem Ziel vorführen.

1.5 Kette 4: Prosumer-, PV- und Direktvermarkter-APIs

Portale für Anlagenregistrierung, Einspeise- und Vergütungsdaten haben mehrere Rollen (Anlagenbetreiber, Installateur, Direktvermarkter, Mitarbeiter), mehrere Mandanten und API-Schlüssel mit Scopes. Ein Fehler in der Mandantentrennung erlaubt, fremde Anlagendaten zu lesen oder vergütungsrelevante Felder wie Leistung oder Inbetriebnahmedatum zu ändern. Greift keine nachgelagerte Prüfung, folgen falsche Vergütungszahlungen, die sich nur mühsam zurückfordern lassen, und offengelegte Anlagendaten, die für Betreiber und Direktvermarkter Geschäftsgeheimnisse sind. Der Test liest mit der Identität eines anderen Mandanten, prüft, ob ein Installateur Zuordnungen ändern darf, die dem Betreiber vorbehalten sind, und ob ein API-Schlüssel über seinen Scope hinaus wirkt.

1.6 Kette 5: Steuerbare Verbraucher nach § 14a EnWG

Die Anmeldung von Wärmepumpen, privaten Ladepunkten oder Speichern als steuerbare Verbrauchseinrichtung ist an reduzierte Netzentgelte geknüpft (§ 14a Abs. 2 EnWG, Festlegung BK8-22/010-A). Wer über ein Installateurs- oder Kundenportal ein nicht existierendes oder mehrfach angemeldetes Gerät anmelden kann, erhält den Vorteil, ohne dass er ihm zusteht. Der Netzbetreiber verliert Netzentgelte und plant mit steuerbarer Last, die es nicht gibt. Der Pentest betrifft die Portal- und API-Schicht vor dem Steuerpfad; das Smart-Meter-Gateway hat eigene BSI-Vorgaben. Steuerfunktionen werden nur am Berechtigungsmodell und an Testkomponenten geprüft, nie im Produktivnetz.

Was die fünf Ketten gemeinsam haben: Sie hängen an Eigenschaften, die sich mit jedem Release ändern können. Die Fehlerklasse entsteht zwischen zwei Prüfterminen laufend neu. Die Frage je Anwendung lautet deshalb: Wie viel Zeit liegt zwischen einer Änderung am Berechtigungsmodell und dem Test, der sie erneut bewertet?

Kapitel 2: Das Programm, Stufung und Takt für ein ganzes Portfolio

2.1 Von der Budgetfrage zur Risikofrage

Ein Versorger betreibt schnell Dutzende Webanwendungen und APIs: Portale, Apps, Schnittstellen für Partner und Dienstleister. Die Frage "wie viele Pentests können wir uns leisten" führt zu Budgetdiskussionen ohne Ergebnis. Tragfähiger ist: Welche Anwendung darf mit welchem Abstand zwischen Änderung und Prüfung leben? Daraus ergibt sich je Anwendung eine Stufe aus zwei Größen.

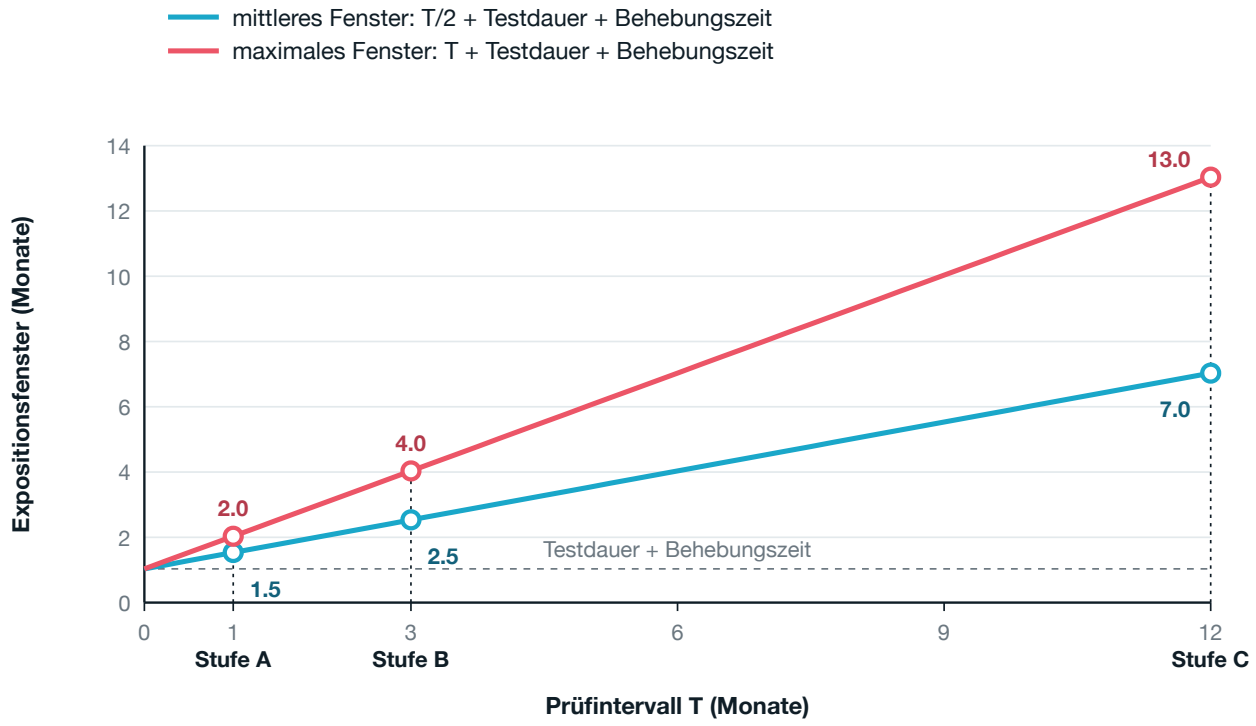
Änderungsrate, gemessen in Releases pro Jahr aus den Daten der Auslieferungskette, nicht aus Schätzungen. Maßgeblich ist, wie oft sich etwas ändert, das das Berechtigungsmodell berühren kann: neue Felder, Endpunkte, Rollen, Integrationen.

Schadenswirkung, bestimmt aus Eigenschaften ohne Messung: Datenklasse (Verbrauchs- und Bankdaten), Geldbezug (Abrechnung, Vergütung, Abschläge), Steuerungsnähe (Ladepunkte, steuerbare Verbraucher), Zahl der Mandanten, Erreichbarkeit aus dem Internet und Rolle im Regelwerk (Kapitel 4). Drei Stufen je Eigenschaft reichen, solange die Rubrik für alle Anwendungen gleich gilt. Für den Bestand hilft ein Attack-Surface-Scan, der die von außen erreichbaren Hosts und Subdomains auflistet; für einen ersten Stufenvorschlag mit Begründung je Anwendung eignet sich ein Planungswerkzeug wie der Pentest Program Planner von SQUR (**squr.ai/plan**), der auch ergeben kann, dass eine Anwendung kein Pentest-Ziel ist.

2.2 Vom Raster zum Takt

Für jede Stufe legen Sie ein **maximal tolerierbares Zeitfenster** zwischen Änderung und Prüfung fest; daraus folgt das Prüfintervall T. Das mittlere Expositionsfenster einer neuen Schwachstelle ist dann T/2 plus Testdauer plus Behebungszeit, vorausgesetzt, der nächste Lauf findet sie. Beim jährlichen Pentest sind das ohne Testdauer und Behebung im Mittel sechs Monate, beim vierteljährlichen sechs Wochen. Bei einem Autonomen Pentest nach jedem Release hängt das Fenster nur noch an Testdauer und Behebungszeit (Abbildung 5, Anhang B).

Damit verschiebt sich der Engpass. Bei monatlichem Takt und einem Monat Behebungszeit macht die Behebung rund zwei Drittel des mittleren Fensters aus, bei einem Test nach jedem Release neun Zehntel. Wer den Takt erhöht, sollte die Behebung deshalb gleich mitplanen: Fristen je Schweregrad, einen Verantwortlichen je Anwendung, einen Retest gleich nach der Korrektur und Korrekturvorschläge, die direkt im Entwicklungsprozess ankommen und zunehmend automatisiert entstehen.



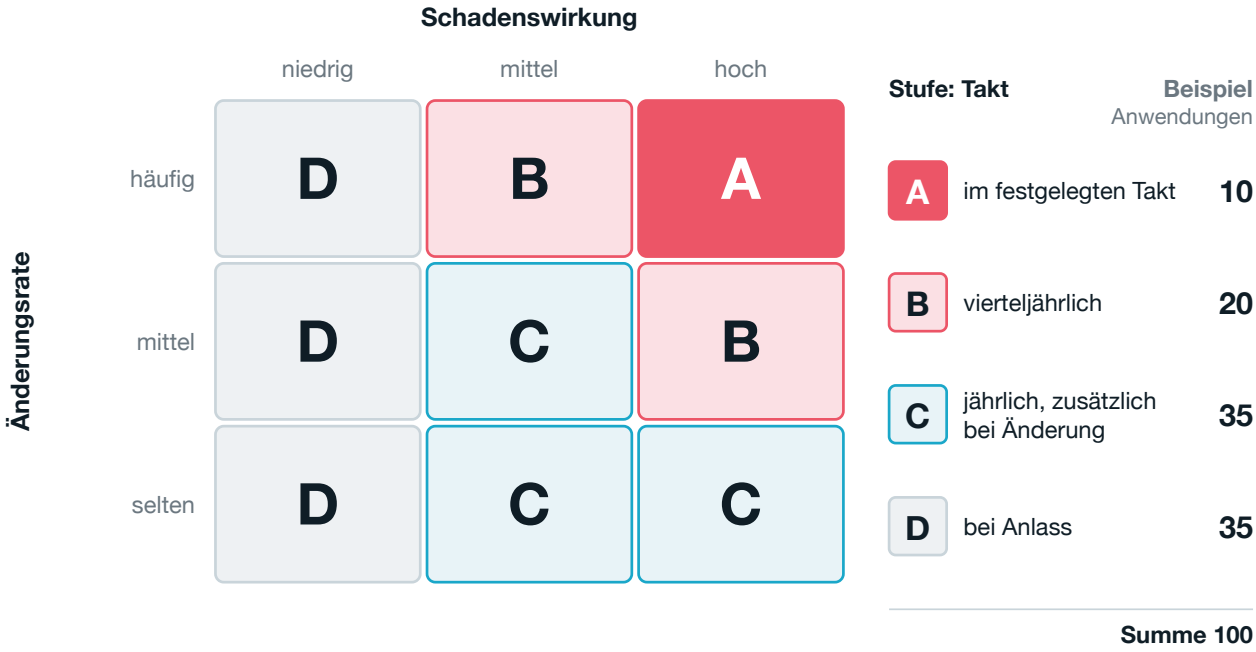
Beispielwerte, schematisch: Testdauer 24 Stunden, Behebungszeit 1 Monat. Markiert: Intervalle 1, 3 und 12 Monate.

Abbildung 5: Erwartetes Expositionsfenster in Abhängigkeit vom Prüfintervall T, schematisch, nach T/2 + Testdauer + Behebungszeit (Mittel) und T + Testdauer + Behebungszeit (Maximum). Beispielwerte: Testdauer 24 Stunden, Behebungszeit 1 Monat; gleiche Erkennung in jedem Lauf angenommen.

Stufe	Änderungsrate x Wirkung	Takt	Umfang und Tiefe	Nachweis
A	häufig x hoch (Kundenportal mit Abrechnung, Ladepunkt-Backend mit Zahlungsbezug, Aggregator-Plattform)	nach jedem größeren Release, mindestens monatlich	mehrere Rollen, Matrix Rolle x Funktion x Objekt, Zustandsübergänge, Retest	Beweis, Retest, Zeitreihe über alle Läufe
B	mittel x hoch oder häufig x mittel (Installateursportal, Einspeise-API)	vierteljährlich	mindestens zwei Rollen, Schwerpunkt auf Änderungen seit dem letzten Lauf	Beweis, Retest, Vergleich zum Vorlauf
C	selten x hoch oder mittel x mittel (stabiler Abrechnungskern, interne Fachanwendung mit Internetzugang)	jährlich und bei wesentlicher Änderung	Authentisierung, Autorisierung und Eingabeverarbeitung entlang der Kernfunktionen	Pentest-Bericht, Nachweis der Auslöserprüfung
D	niedrige Wirkung (statische Seiten, Informationsangebote ohne Anmeldung)	bei Anlass	Überwachung der Außenfläche	Bestandsliste, Prüfung bei Änderung

Die Stufe folgt aus Änderungsrate und Wirkung, nicht aus dem Betriebsmodell. Wird eine Anwendung von Dritten betrieben, ändert das die Wirkung nicht, sondern den Weg zum Nachweis (Abschnitt 2.5).

Neben der Uhr lösen drei Ereignisse einen Lauf aus: eine **wesentliche Änderung** (neue Rolle, neue Schnittstelle, neuer Zahlungs- oder Steuerpfad), ein **Vorfall im Umfeld** und eine **neue Angriffsform**, die auf die Anwendung zutrifft. Die Definition gehört in Ihren Change-Prozess.



Beispielverteilung von 100 Anwendungen und APIs, angenommen, keine Messung. Stufen wie in der Tabelle in Kapitel 2.

Abbildung 6: Stufungsraster aus Änderungsrate und Schadenswirkung mit den Stufen A bis D. Die Verteilung von 100 Anwendungen ist ein angenommenes Beispiel.

2.3 Ein Rechenbeispiel

Angenommen, ein Versorger betreibt 100 Webanwendungen und APIs: 10 in Stufe A, 20 in Stufe B, 35 in Stufe C und 35 in Stufe D. Dann ergeben sich 120 Läufe in Stufe A (zwölf je Anwendung), 80 in Stufe B (vier je Anwendung) und 35 in Stufe C, zusammen rund 235 Läufe im Jahr, zuzüglich der Auslöser. Mit menschlichen Pentests, die je Termin Wochen Vorlauf und mehrere Personentage brauchen, ist das für die meisten Versorger kaum zu organisieren. Autonome Pentests machen denselben Takt praktikabel, weil der Aufwand je Lauf nach der Einrichtung bei Ihnen fast nur noch im Sichten der Ergebnisse liegt. Red-Team-Übungen bleiben frei für die Gesamtprüfung von Angriff und Verteidigung über die ganze Organisation.

2.4 Aufwand und Betrieb

Ob ein Programm trägt, entscheidet der laufende Aufwand. Er verteilt sich so:

	Einmalig je Anwendung	Laufend
Sie	Freigabe des Betreibers, Freigabeliste, Testidentitäten und Testobjekte, Abbruchweg, fachliche Regeln, die sich nicht von selbst ergeben	Konten und Testdaten pflegen, wenn sich Rollen ändern; Befunde sichten und zuweisen; Behebung terminieren
Anbieter	Einrichtung des Ziels, Abnahmetests aus Kapitel 3	Läufe im vereinbarten Takt, Bericht mit Beweis, Retest nach Korrektur, Meldung, wenn eine Anmeldung scheitert oder ein Bereich nicht erreicht wurde

Zwei Punkte entscheiden, ob der laufende Teil klein bleibt. Erstens: Ein Lauf, bei dem die Anmeldung scheitert, darf nicht als "keine Befunde" erscheinen, sondern muss als unvollständig gemeldet werden, mit Angabe der nicht erreichten Bereiche. Zweitens: Die Ergebnisse müssen so belegt sein, dass das Sichten Minuten statt Stunden dauert, also mit reproduzierbarer Anfrage, Gegenprobe und klarer Einordnung. Lassen Sie sich beides an einem Beispielbericht zeigen.

2.5 Dienstleister und Konzernverbünde

Betreibt Konzern-IT oder ein Dienstleister die Anwendung, muss die Freigabe entlang der Kette schriftlich geklärt sein; ohne Befugnis des Betreibers ist ein Test nicht gedeckt (§ 202a StGB). Bei fremder Software-as-a-Service gehört die Sicherheit der Lieferkette zu Ihren Pflichten (§ 30 Abs. 2 Nr. 4 BSIG): Prüfen Sie Nachweise des Anbieters nach Umfang, Qualität und Aktualität, vereinbaren Sie Testrechte, wo die Wirkung hoch ist, und testen Sie selbst, was Sie testen dürfen. In Konzernverbünden hilft eine Portfolio-Sicht über alle Gesellschaften, ein einheitliches Raster und eine Trennung der Ergebnisse: Jede Gesellschaft sieht ihre Schwachstellen, die Konzernsicht die Kennzahlen.

Kapitel 3: Was Sie von einem Anbieter Autonomer Pentests verlangen sollten

3.1 Der Grundsatz

Ein Autonomer Pentest ist empfehlenswert, wenn seine Grenzen außerhalb des Agenten durchgesetzt werden: Der Agent schlägt Aktionen vor, eine getrennte Prüfinstanz entscheidet über jede einzelne, und Sie können jederzeit abbrechen. Das ist die Voraussetzung für einen sicheren Betrieb; die Qualität der Erkennung zeigt erst ein Lauf an Ihrer eigenen Anwendung. Als gemeinsames Vokabular eignet sich die BSI-Studie "Durchführungskonzept für Penetrationstests" mit vier Stufen der Aggressivität (passiv, vorsichtig, abwägend, aggressiv); der Anbieter sollte seine Betriebsarten darauf abbilden können.

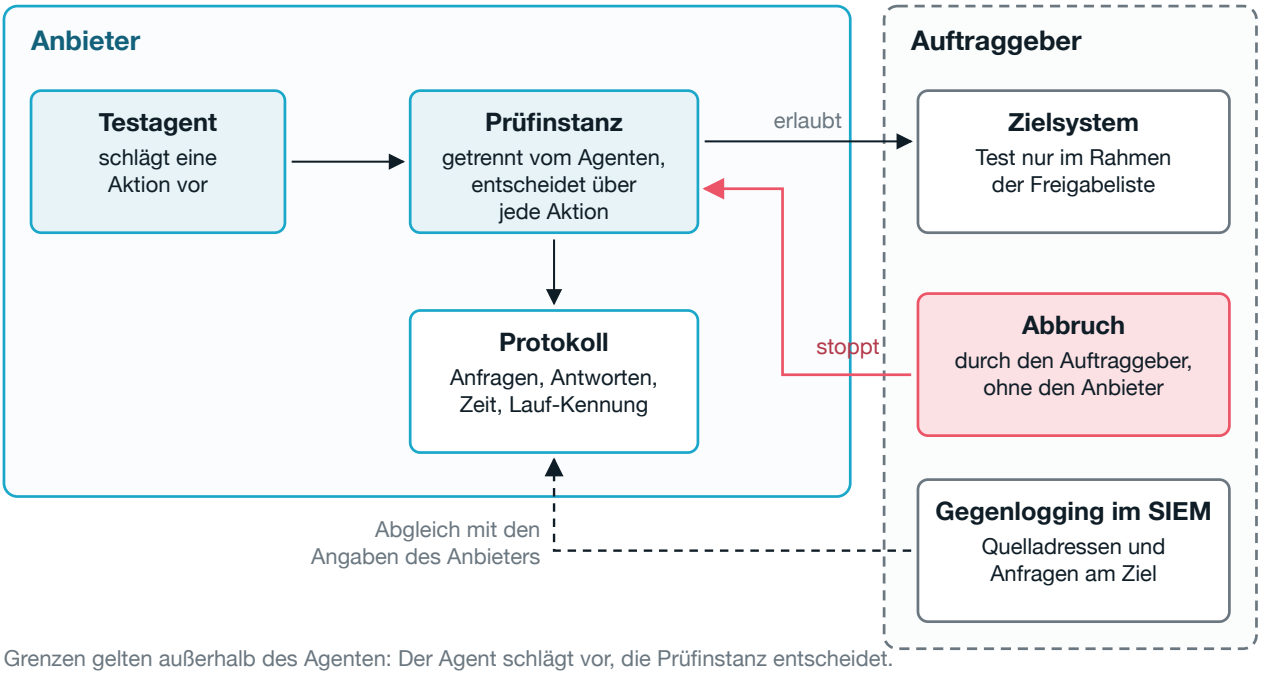


Abbildung 7: Anforderungsschichten. Der Testagent schlägt vor, eine getrennte Prüfinstanz entscheidet, der Auftraggeber kann abbrechen und gleicht seine eigenen Protokolle mit den Angaben des Anbieters ab.

3.2 Sieben Anforderungen

1. **Scope-Durchsetzung.** Eine Freigabeliste mit Host, Port, Pfad und Betriebsart je Adresse, die vor jeder Aktion geprüft wird, auch nach Weiterleitungen.
2. **Begrenzte Eingriffstiefe.** Zerstörerische Handlungen sind schriftlich ausgeschlossen; je Adresse ist festgelegt, ob nur bestätigt oder auch ausgenutzt wird. Schreibende Tests treffen nur Ihre Testobjekte.
3. **Abbruch.** Sie können einen laufenden Test selbst abbrechen, ohne den Anbieter zu kontaktieren; der Anbieter kann es ebenfalls.
4. **Nachvollziehbarkeit.** Jede Schwachstelle steht mit reproduzierbarem Beweis im Bericht; der Anbieter beschreibt, was protokolliert wird und wie das Protokoll geschützt ist.
5. **Verbotene Handlungen.** Keine Ansprache von Personen, kein Ausleiten von Daten über den minimalen Beweis hinaus, keine Wiederverwendung gefundener Zugangsdaten außerhalb des

Scopes, keine Persistenz. Je Punkt: technisch unterbunden oder durch Vorgaben gesteuert.

6. **Aufbewahrung und Datenschutz.** Fristen für Protokolle, Beweise und Zugangsdaten, eine Schwärzungsregel, Auftragsverarbeitung und Verarbeitungsorte im Vertrag.
7. **Einsatz von Sprachmodellen.** Ein Sprachmodell darf eine Freigabe höchstens verschärfen, nie erweitern, und ein regelbasierter Boden bleibt bei Ausfall bestehen. Verhaltensänderungen werden vorab angekündigt.

3.3 So prüfen Sie es selbst

Eine Zusage, die nur der Anbieter belegen kann, ist keine Kontrolle. Drei Abnahmetests liegen bei Ihnen. Vor dem ersten Lauf genügt es, Freigabeliste und Abbruchweg gemeinsam mit dem Anbieter zu prüfen; die formale Abnahme führen Sie durch, bevor Sie Produktivsysteme oder das ganze Portfolio in den Takt nehmen. Jeder prüft eine bestimmte Kontrolle, nicht die Qualität der Erkennung:

- **Der Köder prüft die Scope-Durchsetzung.** Tragen Sie einen Host, dessen Verbindungen Sie protokollieren, als ausdrücklich verboten ein und setzen Sie auf einer getesteten Seite einen Verweis auf ihn. Bei korrekter Durchsetzung bleibt Ihr Protokoll leer, und das Protokoll des Anbieters zeigt den Versuch als blockiert. Ein leeres Protokoll ohne blockierten Versuch belegt nichts.
- **Die Stoppmessung prüft den Abbruch.** Lösen Sie im Lauf Ihren Abbruch aus und messen Sie, wann keine Anfrage mehr ankommt.
- **Das Gegenlogging prüft die Nachvollziehbarkeit.** Gleichen Sie in Ihrem SIEM oder Webserver-Protokoll ab, dass alle Anfragen aus den genannten Quelladressen und Zeitfenstern kommen.

Wer diese drei Tests besteht, hat gezeigt, dass seine Grenzen nicht nur auf dem Papier stehen.

Der erste Lauf: So steigen Sie ein

Ein Programm beginnt mit einer Anwendung. Wählen Sie eine, die oft geändert wird und deren Fehler wehtun würden, typischerweise das Kundenportal oder seine API.

Was Sie vorbereiten: einen Verantwortlichen für die Anwendung und ihre Befunde, die Freigabe des Betreibers, die Testumgebung mit Freigabeliste und geprüftem Abbruchweg, zwei Testkunden und ein Mitarbeiterkonto mit Testverträgen. Erfahrungsgemäß ist das in wenigen Tagen erledigt.

Was Sie bekommen: einen vollständigen Lauf mit Bericht, Beweis je Schwachstelle und Angabe der nicht erreichten Bereiche, einen Retest nach Ihrer Korrektur und eine Einschätzung, welche weiteren Anwendungen in welchen Takt gehören.

Was Sie danach entscheiden: ob und wie Sie das Programm auf weitere Anwendungen ausweiten. Vor diesem Schritt stehen die Abnahmetests aus Abschnitt 3.3.

Die ersten 30 Tage

- **Woche 1, Vorbereitung und erster Lauf.** Verantwortlichen, Freigabe, Testumgebung und Testidentitäten für die erste Anwendung klären und festlegen, wie mit Befunden umgegangen wird. Dann autonom testen lassen.

- **Woche 2, Inventar.** Webanwendungen und APIs aus Pipelines, DNS, Zertifikaten und Dienstleisterverträgen, je mit Betreiber. Einen schnellen Startpunkt für die von außen sichtbaren Hosts und Subdomains liefert ein Attack-Surface-Scan; der von SQUR ist frei nutzbar und braucht kein Konto (**squr.ai/attack-surface**).
- **Woche 3, Raster, Abnahme und Ausbau.** Änderungsrate aus den Deployment-Protokollen, Stufen A bis D vergeben; einen ersten Vorschlag je Anwendung mit Takt, Methode und Begründung liefert der Pentest Program Planner, ebenfalls frei nutzbar ohne Konto (**squr.ai/plan**). Köder, Stoppmessung und Gegenlogging durchführen, dann Stufe A in den Takt aufnehmen.
- **Woche 4, Behebung und Dokument.** Fristen je Schweregrad festlegen, Stufung, Takt und Begründung auf zwei Seiten dokumentieren, Regimespalte bestätigen lassen.

Den ersten Lauf richten wir mit Ihnen in einem kurzen Gespräch ein: squr.ai/meeting. Im Erstgespräch klären wir Eignung, Umfang und Vorbereitung; Konditionen und Leistungen des ersten Tests vereinbaren wir, bevor Sie ihn beauftragen. Der QR-Code führt direkt zur Terminwahl.

Schluss: Das Programm in einem Absatz

Ein Pentest-Programm besteht aus drei schriftlichen Entscheidungen. Der **Takt** ergibt sich aus der Änderungsrate und dem tolerierten Zeitfenster. Die **Tiefe** ergibt sich aus der Schadenswirkung: Wo Berechtigungslogik Geld, Mandanten oder Steuerung berührt, braucht der Test mehrere Identitäten und fachliche Bewertung. Der **Nachweis** ergibt sich aus dem Bericht: Schwachstelle, Beweis, Frist, Retest. Autonome Pentests machen den Takt für das ganze Portfolio praktikabel, und mit den Abnahmetests aus Kapitel 3 prüfen Sie selbst, dass ihre Grenzen gelten. Beginnen Sie mit der ersten Anwendung, bevor Inventar und Raster fertig sind; erweitern lässt sich das Programm jederzeit.

Die folgenden Kapitel 4 und 5 sind Referenz: wie das Programm die gesetzlichen Pflichten belegt und an Rahmenwerke anschließt.

Kapitel 4: Der Nachweis gegenüber Aufsicht und Geschäftsleitung

4.1 Was das Recht verlangt

Alle Angaben beruhen auf Normtexten, abgerufen am 4. und 8. November 2026. Die Einordnung ist eine Lesart des Wortlauts und keine Rechtsauskunft.

Das Gesetz zur Umsetzung der NIS-2-Richtlinie (BGBl. 2025 I Nr. 301) ist seit dem 6. Dezember 2025 in Kraft. § 30 Abs. 1 BSIG verlangt geeignete, verhältnismäßige und wirksame Maßnahmen, und ihre Einhaltung ist zu dokumentieren. § 30 Abs. 2 nennt ausdrücklich das Management von Schwachstellen (Nr. 5) und Verfahren zur Bewertung der Wirksamkeit (Nr. 6). Für Netz- und Anlagenbetreiber verlangt § 5d EnWG die Dokumentation der Einhaltung gegenüber der Bundesnetzagentur. Einen bestimmten Prüftakt schreibt das Gesetz nicht vor; Sie legen ihn selbst fest und begründen ihn aus Ihrer

Risikoanalyse. Genau dafür ist die Stufung aus Kapitel 2 gemacht. Ein Pentest-Programm mit Retest liefert dazu konkrete technische Belege für die geprüften Anwendungen und Kontrollen; die übrigen Pflichten deckt es nicht ab.

Gegenstand	Bis 5. Dezember 2025	Seit 6. Dezember 2025
Rechtsgrundlage	IT-SiG 2.0 mit § 8a BSIG alt; für Energie § 11 Abs. 1a und 1b EnWG alt	NIS2UmsuCG mit neu gefasstem BSIG; §§ 5c bis 5e EnWG
Adressaten	Betreiber Kritischer Infrastrukturen	Besonders wichtige und wichtige Einrichtungen (§ 28 BSIG); Anlage 1 nennt ausdrücklich Strom- und Gaslieferanten, Ladepunktbetreiber und Aggregatoren
Dokumentation	Nachweis alle zwei Jahre (nicht für Energienetz- und Anlagenbetreiber)	Pflicht für alle, die Einhaltung zu dokumentieren (§ 30 Abs. 1 Satz 3 BSIG, § 5d Abs. 1 EnWG)
Geschäftsleitung	im geprüften Material nicht geregelt	Umsetzung und Überwachung, Haftung, Schulung (§ 38 BSIG, § 5e EnWG)

Neu ist damit vor allem zweierlei: Vertrieb und Ladeinfrastruktur sind eigene Einrichtungsarten, und Ihre Portale und APIs liegen genau dort. Und die Geschäftsleitung überwacht persönlich.

4.2 Welches Regime für welche Anwendung

Für Betreiber von Energieversorgungsnetzen, Energieanlagen und digitalen Energiediensten, die den §§ 5c bis 5e EnWG unterliegen, gelten die §§ 30, 31, 32, 35, 36, 38, 39, 61 und 62 BSIG nicht (§ 28 Abs. 5 BSIG). Das gilt nicht, soweit die Einrichtung weitere kritische Anlagen betreibt oder durch weitere Tätigkeiten einer anderen Einrichtungsart zuzuordnen ist. Ein Stadtwerk mit Netz, Vertrieb und Ladeinfrastruktur kann daher je Tätigkeit unter unterschiedlichen Regimen stehen. Praktisch erhält jede Anwendung neben ihrer Stufe eine **Regimespalte**, die Ihre Rechtsabteilung bestätigt. Kundenportale und APIs liegen überwiegend im Vertrieb und der Ladeinfrastruktur und damit eher im BSIG-Regime.

Anwendungstyp	EnWG §§ 5c bis 5e (Katalog der BNetzA)	BSIG §§ 30 ff.
Netzbetrieb Systeme für den sicheren Netzbetrieb	●	—
Vertriebsportal Stromlieferant, Anlage 1 Nr. 1.1.1	—	●
Ladeinfrastruktur Ladepunktbetreiber, Anlage 1 Nr. 1.1.9	—	●
Aggregator, steuerbare Einrichtungen Portale berühren Steuerpfade	◌	◌

● nach Wortlaut eher einschlägig ◌ Grenze offen: Abgrenzung ausdrücklich klären und dokumentieren

Lesart des Wortlauts mit Stand 4. November 2026, keine Rechtsauskunft. Welches Regime für Ihre Gesellschaft und Ihre Tätigkeiten gilt, ist eine Rechtsfrage; dasselbe Stadtwerk kann je Tätigkeit unter unterschiedlichen Regimen stehen.

Abbildung 8: Regimekarte, Lesart des Wortlauts mit Stand 8. November 2026, keine Rechtsauskunft.

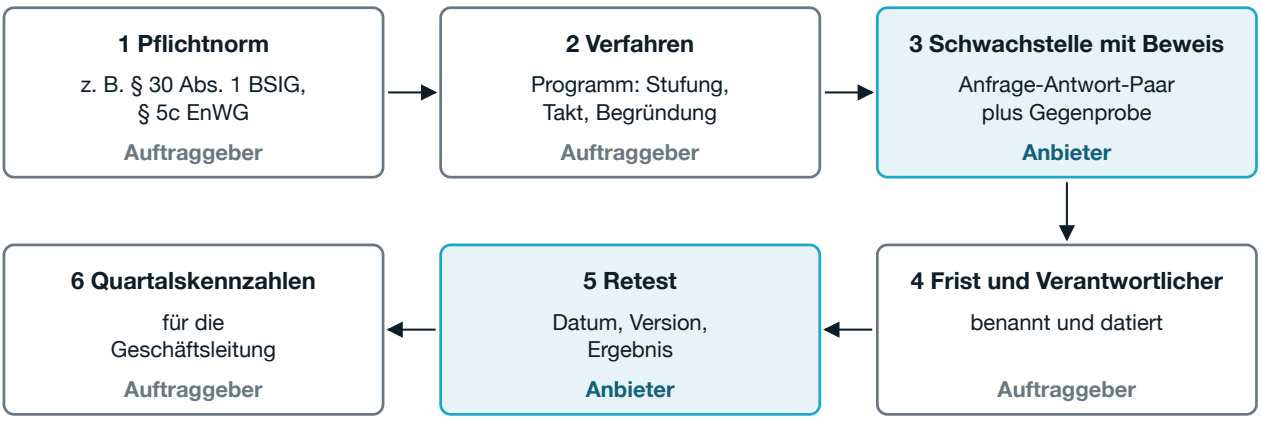
4.3 Wie Pentest-Ergebnisse die Pflichten belegen

Norm	Inhalt nach Wortlaut	Beitrag des Pentest-Programms
§ 30 Abs. 1 BSIG	wirksame Maßnahmen, Einhaltung dokumentiert	Programmdokument mit Stufung, Takt und Begründung; Berichte
§ 30 Abs. 2 Nr. 5 BSIG	Management von Schwachstellen	direkt: Schwachstellen, Behebung und Retest
§ 30 Abs. 2 Nr. 6 BSIG	Bewertung der Wirksamkeit	direkt: der Test ist ein solches Verfahren, die Zeitreihe zeigt die Entwicklung
§ 30 Abs. 2 Nr. 9 und 10 BSIG	Zugriffskontrolle, Multi-Faktor-Authentisierung	Belege zu Autorisierung und Anmeldung
§§ 5c, 5d EnWG	angemessener Schutz, Dokumentation, Mängelbeseitigung in Frist	Schwachstelle, Frist, Verantwortlicher, Nachweis der Beseitigung
§ 38 BSIG, § 5e EnWG	Umsetzung und Überwachung durch die Geschäftsleitung	Quartalsbericht mit Kennzahlen als ein Baustein
§§ 39, 61, 62 BSIG	Audits, Prüfungen, anlassbezogene Aufsicht	laufend geführte Berichte, die einem Dritten ohne Erläuterung verständlich sind

Das KRITIS-Dachgesetz (in Kraft seit 17. März 2026) regelt die physische Resilienz; für Software und IT-Dienste, die nicht unmittelbar der Steuerung, Überwachung oder Unterstützung physischer Prozesse dienen, gelten ausschließlich die Vorgaben des BSIG (§ 2 Nr. 2 KRITISDachG).

4.4 Der Bericht, den Prüfer verstehen

Die Aufsicht kann Dokumentation, Mängelbeseitigungsplan und Nachweis der Beseitigung verlangen (§ 5d EnWG, §§ 39 und 61 BSIG). Daraus folgt ein Befund mit fünf Feldern: **Schwachstelle** und **Beweis** (Anfrage-Antwort-Paar mit Gegenprobe) vom Anbieter, **Frist und Verantwortlicher** sowie **Behebung** von Ihnen, **Retest-Ergebnis** mit Datum wieder vom Anbieter. Lesbar für Dritte wird das durch eine beständige Kennung je Schwachstelle über alle Läufe.



Beispieldatensatz einer Schwachstelle (Felder, ohne reale Inhalte)

Kennung	Fehlerklasse	Beweis	Frist	Verantwortlicher	Retest
V-0001	Zuordnung nicht geprüft	Anfrage, Antwort, Gegenprobe	Datum	Rolle	Datum, Version, Ergebnis

Blau hinterlegt: liefert der Anbieter. Weiß: legt der Auftraggeber fest oder verantwortet er.
Jede Schwachstelle behält ihre Kennung über alle Läufe; Beweis und Bewertung stehen getrennt.

Abbildung 9: Nachweiskette von der Pflichtnorm bis zu den Quartalskennzahlen, mit Beispieldatensatz einer Schwachstelle ohne reale Inhalte.

Für die Geschäftsleitung eignet sich ein Quartalsbericht mit vier Kennzahlen: Abdeckung je Stufe (Anteil der im vereinbarten Takt geprüften Anwendungen), offene Schwachstellen nach Schweregrad und Alter, mittlere Zeit bis zur Behebung und Retest-Quote. Er unterstützt die Überwachung des Pentest-Programms und ist ein Baustein für den Nachweis, dass die Geschäftsleitung die Maßnahmen nach § 30 BSIG umsetzt und überwacht.

4.5 Zeitachse

- **6. Dezember 2025:** NIS2-Umsetzungsgesetz mit neu gefasstem BSIG in Kraft.
- **17. März 2026:** KRITIS-Dachgesetz in Kraft.
- **Herbst 2026, angekündigt:** Konsultation des IT-Sicherheitskatalogs für digitale Energiedienste; bis zu neuen Katalogen gelten die Kataloge für Netze (2015) und Anlagen (2018) fort.
- **Frühestens 6. Dezember 2028:** allgemeine Nachweisanforderung für besonders wichtige Einrichtungen (§ 61 Abs. 3 BSIG). Einzelne Audits kann das BSI schon jetzt anordnen.

Kapitel 5: Anschluss an Rahmenwerke

Die gängigen Rahmenwerke verlangen regelmäßige Prüfungen, nennen aber kein Intervall. Das Programm liefert die Belege für die Teile, die von Anwendungen und Schnittstellen abhängen, und ersetzt keines davon.

- **B3S Aggregatoren (BDEW, Version 1.2), A 3.7.5:** "Durch regelmäßige interne Audits und technische Sicherheitstests muss überprüft werden, dass die Anforderungen dieses B3S eingehalten werden."

- **B3S Fernwärme (BDEW und AGFW, Version 1.2), 4.9.2:** Die Wirksamkeit der Maßnahmen "muss regelmäßig in geeigneter Weise überprüft werden", beispielhaft anlassbezogen nach Änderungen an IT- oder Kommunikationssystemen. Das entspricht den Auslösern aus Kapitel 2.
- **IT-Sicherheitskatalog der BNetzA (Strom und Gas, 2015):** verlangt ein zertifiziertes ISMS nach ISO/IEC 27001 und 27019. Der Pentest belegt die Wirksamkeit technischer Maßnahmen, das Zertifikat bleibt eigenständig.
- **IEC 62443:** Wo Portale und APIs an Steuerpfade grenzen (Ladepunkte, § 14a, Aggregatoren), prüft das Programm diese Grenze von der IT-Seite: Rechte, Trennung und ob sich über ein Portal oder eine API ein Befehl in Richtung Steuerung auslösen lässt. Wie sich das in eine Zonierung nach IEC 62443 einfügt, klären Sie mit Ihrer OT-Sicherheit.

Anhang A: Checkliste zum Mitnehmen

Für jeden Anbieter. Die Verweise beziehen sich auf den Fachbeitrag "Autonome Pentests für Energieversorger: Sicherheit im Takt Ihrer Releases".

1. Erste Anwendung und Bestand (Kapitel 2)

- ☐ Für die erste Anwendung sind Verantwortlicher, Freigabe des Betreibers, Testumgebung, Testidentitäten und der Umgang mit Befunden geklärt; der erste Lauf ist terminiert, bevor Inventar und Stufung fertig sind.
- ☐ Alle Webanwendungen und APIs sind inventarisiert, auch die von Konzern-IT und Dienstleistern betriebenen, mit benanntem Betreiber; die von außen sichtbaren Hosts sind per Attack-Surface-Scan abgeglichen.
- ☐ Je Anwendung sind Änderungsrate (Releases je Jahr aus den Pipeline-Daten) und Schadenswirkung (Datenklasse, Geld, Steuerungsnähe, Mandanten, Erreichbarkeit, Regime) festgehalten.
- ☐ Jede Anwendung hat eine Stufe A bis D und ein maximal tolerierbares Zeitfenster zwischen Änderung und Prüfung. Für fremdbetriebene Anwendungen ist der Weg zum Nachweis festgelegt.
- ☐ Auslöser sind im Change-Prozess verankert: wesentliche Änderung, Vorfall im Umfeld, neue Angriffsform.

2. Befugnis und Regeln (Kapitel 3)

- ☐ Der Betreiber jedes Zielsystems hat den Test schriftlich freigegeben.
- ☐ Freigabeliste mit Host, Port und Betriebsart je Adresse; zerstörerische Handlungen sind ausgeschlossen; die Betriebsarten des Anbieters sind auf die vier Stufen des BSI-Rasters abgebildet.
- ☐ Zwei bis drei Testidentitäten je Rolle und eigene Testobjekte sind bereitgestellt; schreibende Tests und Steuerbefehle gelten nur dort, nie im Produktivbetrieb.
- ☐ Verbotene Handlungen stehen im Vertrag: Ansprache von Personen, Datenabfluss über den minimalen Beweis hinaus, Wiederverwendung von Zugangsdaten, Persistenz.

3. Abnahmetest vor dem Ausbau auf Produktivsysteme (Abschnitt 3.3)

- ☐ Köder-Host (verboten eingetragen, jede Verbindung protokolliert): eigenes Protokoll leer, Protokoll des Anbieters zeigt den blockierten Versuch.
- ☐ Stoppmessung für beide Abbruchwege; beide Zeiten stehen im Abnahmeprotokoll.
- ☐ Gegenlogging: Quelladressen und Zeitfenster des Anbieters stimmen mit dem eigenen Protokoll überein; Abweichungen sind erklärt.
- ☐ Ein Lauf mit gescheiterter Anmeldung wird als unvollständig gemeldet, nicht als "keine Befunde".
- ☐ Aufbewahrungs- und Löschfristen, Schwärzungsregel und Verarbeitungsorte stehen im Vertrag.

4. Behebung und Nachweis (Kapitel 4)

- ☐ Jeder Befund trägt fünf Felder: Schwachstelle, Beweis (Anfrage-Antwort-Paar mit Gegenprobe), Frist und Verantwortlicher, Behebung, Retest-Ergebnis. Die Kennung bleibt über alle Läufe gleich.
- ☐ Für die Behebung gelten Fristen je Schweregrad und ein Verantwortlicher je Anwendung; der Retest folgt direkt auf die Korrektur.
- ☐ Ein Quartalsbericht mit vier Kennzahlen liegt der Geschäftsleitung vor: Abdeckung je Stufe, offene Schwachstellen nach Schwere und Alter, mittlere Behebungszeit, Retest-Quote.
- ☐ Jede Anwendung hat eine Regimespalte (BSIG, § 5c EnWG, nicht erfasst), von der Rechtsabteilung bestätigt.
- ☐ Ein Programmdokument mit Stufung, Takt und Begründung liegt vor, bevor jemand danach fragt (§ 30 Abs. 1 Satz 3 BSIG, § 5d EnWG).

Rechtsstand 8. November 2026. Keine Rechtsauskunft. Ersten Lauf vereinbaren: ***squr.ai/meeting***

Anhang B: Rechenblatt Expositionsfenster

Das Rechenblatt `Rechenblatt-Expositionsfenster.xlsx` (nur Formeln, keine Makros) rechnet das Expositionsfenster mit Ihren Zahlen; Sie erhalten es zusammen mit diesem Beitrag oder auf Anfrage über **squr.ai/meeting**.

Eingaben. F, Behebungszeit in Tagen vom Befund bis zum bestandenen Retest; **R**, Releases je Jahr, die das Berechtigungsmodell berühren können; **p**, Anteil der Änderungen mit neuer Berechtigungsschwachstelle (kürzt sich beim Vergleich der Verfahren heraus); je Verfahren **T**, Prüfintervall, und **D**, Testdauer, in Tagen. Bei einem Test nach jedem Release tritt an die Stelle von T der Verzug **L** zwischen Release und Test.

Formeln. Mittleres Fenster bei Kalendertakt: T / 2 + D + F; nach jedem Release: L + D + F. Schwachstellen-Tage je Jahr, also die Summe der offenen Tage über alle neuen Schwachstellen: R x p x mittleres Fenster. Überlappen sich mehrere Schwachstellen, ist das mehr als die Zahl der Kalendertage mit mindestens einer offenen Schwachstelle.

Beispiel (Annahmen, keine Messwerte): F = 30 Tage, L = 2 Tage, menschlicher Test D = 14, autonomer Test D = 1. Angenommen ist, dass beide Verfahren die betrachtete Schwachstelle im nächsten Lauf gleichermaßen finden; das Beispiel zeigt also den Effekt von Prüftakt und Testdauer bei gleicher Erkennung und Behebungszeit. Der jährliche Pentest (T = 365) ergibt ein mittleres Fenster von 226,5 Tagen, der Autonome Pentest nach jedem Release 33 Tage, also rund 85 Prozent weniger. Davon sind 30 Tage Behebungszeit: Ist der Takt hergestellt, wird die Behebung zur Stellgröße.

Modell mit echten Release-Terminen. Zwei quelloffene Ladeinfrastruktur-Backends veröffentlichten in den 12 Monaten bis 7. November 2026 7 (SteVe) und 9 (CitrineOS) stabile Releases (GitHub-Releases-API). Mit diesen Release-Terminen und einem angenommenen Jahrestest ergibt das Modell bei jährlichem Test ein mittleres Fenster von 261 bzw. 216 Tagen, bei Test nach jedem Release 33 Tage. Termine, Testdatum und Rechnung stehen im Rechenblatt; es ist ein Modell, keine gemessene Exposition.

Stand der Recherche und Quellen

Rechtsstand 8. November 2026. Die zitierten Normtexte wurden am 4. und 8. November 2026 abgerufen und verglichen. Offen und hier nicht bewertet sind der Stand der Konsultation zum Katalog für digitale Energiedienste, die Verordnung nach dem KRITISDachG und die Normtexte der IEC 62443. Welches Regime für Ihre Gesellschaft gilt, ist eine Rechtsfrage.

- BSIG in der Fassung des NIS2UmsuCG (BGBl. 2025 I Nr. 301, zuletzt geändert durch Art. 8 G v. 22. Oktober 2026, BGBl. 2026 I Nr. 318): **gesetze-im-internet.de/bsig_2025**
- EnWG §§ 5c, 5d, 5e, 14a: **gesetze-im-internet.de/enwg_2005**
- KRITIS-Dachgesetz (BGBl. 2026 I Nr. 66): **gesetze-im-internet.de/kritisdachg**
- StGB § 202a: **gesetze-im-internet.de/stgb/__202a.html**
- Bundesnetzagentur, IT-Sicherheit (Stand 10. August 2026) und IT-Sicherheitskatalog Strom und Gas (August 2015): **bundesnetzagentur.de, IT-Sicherheit**

- Bundesnetzagentur, Festlegung BK8-22/010-A (§ 14a EnWG)
- BDEW, B3S Aggregatoren, Version 1.2: **bdew.de, B3S Aggregatoren**
- BDEW und AGFW, B3S Fernwärme, Version 1.2: **bdew.de, B3S Fernwärme**
- BSI, Durchführungskonzept für Penetrationstests: **bsi.bund.de, Penetrationstest**
- ISA/IEC 62443 Series of Standards: **isa.org, IEC 62443**
- OWASP API Security Top 10 2023: **owasp.org, API Security Top 10**
- OWASP, Authorization Testing Automation Cheat Sheet: **cheatsheetseries.owasp.org**
- Open Charge Alliance, Sicherheits-Whitepaper für OCPP 1.6-J, Edition 4: **openchargealliance.org**
- NIST, National Vulnerability Database, API 2.0, CVEs je Veröffentlichungsmonat ohne zurückgezogene Einträge, Abruf 7. November 2026
- CISA, CSAF-Archiv (**github.com/cisagov/CSAF**), Abruf 6. November 2026
- GitHub-Releases-API der Projekte steve-community/steve und citrineos/citrineos-core, Abruf 8. November 2026

Über SQUR SQUR aus Weingarten (Baden) entwickelt Autonomous Pentesting für Webanwendungen und APIs: authentifiziert in mehreren Rollen, Ergebnis in 24 Stunden, mit Beweis je Schwachstelle und Retest. Für jede Adresse legen Sie fest, ob Schwachstellen nur bestätigt oder auch ausgenutzt werden, und Sie können einen laufenden Test jederzeit selbst abbrechen. Wie Scope-Prüfung, Abbruch und Protokollierung bei SQUR umgesetzt sind, zeigen wir Ihnen gern in einem technischen Review; die Anforderungen aus Kapitel 3 eignen sich dafür als Fragenliste. Gründer, CEO und Autor ist Adam Lundqvist (**LinkedIn**), zuvor Director of Engineering bei Cobalt, dem Unternehmen, das Pentesting-as-a-Service definiert hat. **squr.ai** · Ersten Lauf vereinbaren: **squr.ai/meeting**