

Red-Team-Übung: der kontrollierte Angriff

Es gibt zwei Arten von Häusern: solche, die schon einmal jemand ernsthaft angegriffen hat, und solche, die noch glauben, sie seien in Ordnung. Der Unterschied zwischen beiden ist keine Frage der Technik, sondern eine Frage des Zeitpunkts. Irgendwann kommt jemand vorbei und probiert es aus. Die einzige Entscheidung, die du wirklich treffen kannst, ist die: Möchtest du beim ersten Mal dabei sein und eine Rechnung dafür bekommen — oder lieber nicht dabei sein und eine sehr viel größere Rechnung dafür bekommen.

Eine Red-Team-Übung ist die erste Variante. Ein Team von Menschen, die dafür bezahlt werden, sich wie ein Angreifer zu verhalten, bekommt ein Ziel und ein paar Wochen Zeit. Nicht „findet Schwachstellen“, sondern „kommt an die Konstruktionsdaten“ oder „lest die Postfächer der Geschäftsführung“. Sie dürfen alles, was ein echter Angreifer auch dürfte, abzüglich der Dinge, die im Vertrag ausdrücklich ausgeschlossen sind. Und der Rest deines Hauses weiß nichts davon. Genau das ist der Punkt.

Das klingt für viele nach Hollywood, nach Kapuzenpulli und grünem Text auf schwarzem Grund. In Wirklichkeit ist es zu neunzig Prozent Projektarbeit: Scoping-Workshops, ein Vertrag mit Anlagen, Abstimmungen mit Betriebsrat und Datenschutz, Protokolle mit Zeitstempeln und am Ende ein Bericht, den jemand lesen und in Arbeitspakete übersetzen muss. Die spannenden zehn Prozent sind die, über die man abends redet. Die anderen neunzig entscheiden darüber, ob die Übung dir etwas bringt oder nur ein PDF im Qualitätsmanagement-Ordner erzeugt.

Dieser Beitrag gehört zum Handlungsfeld E der Serie [Microsoft 365 Security](#). Er sortiert zuerst die Begriffe — Pentest, Red Team, Purple Team werden im Angebotsgespräch gern durcheinandergeworfen, und zwar nicht immer versehentlich. Danach geht es darum, was ein Red Team mit dem Schwerpunkt Microsoft 365 tatsächlich anfasst, wie Scope und Rules of Engagement aussehen müssen, warum Betriebsrat und Datenschutz vor und nicht nach der Übung an den Tisch gehören — und was du aus dem Bericht machst, damit er mehr wird als Papier.

FAKTEN — Warum die Beispiele in diesem Beitrag nicht ausgedacht sind

Im Januar 2024 beschrieb Microsoft den Angriff der Gruppe Midnight Blizzard auf die eigene Unternehmensumgebung. Einstieg war ein Password Spray über Wohnadress-Proxys auf ein altes Testkonto ohne MFA in einem Nicht-Produktions-Tenant. Von dort ging es über eine vergessene Legacy-OAuth-Anwendung weiter, die in der Produktivumgebung erhöhte Rechte besaß. Der Angreifer legte eigene Anwendungen an und vergab sich am Ende die Rolle `full_access_as_app` für Exchange Online — also Zugriff auf Postfächer. Genau diesen Pfad läuft ein M365-Red-Team heute nach.

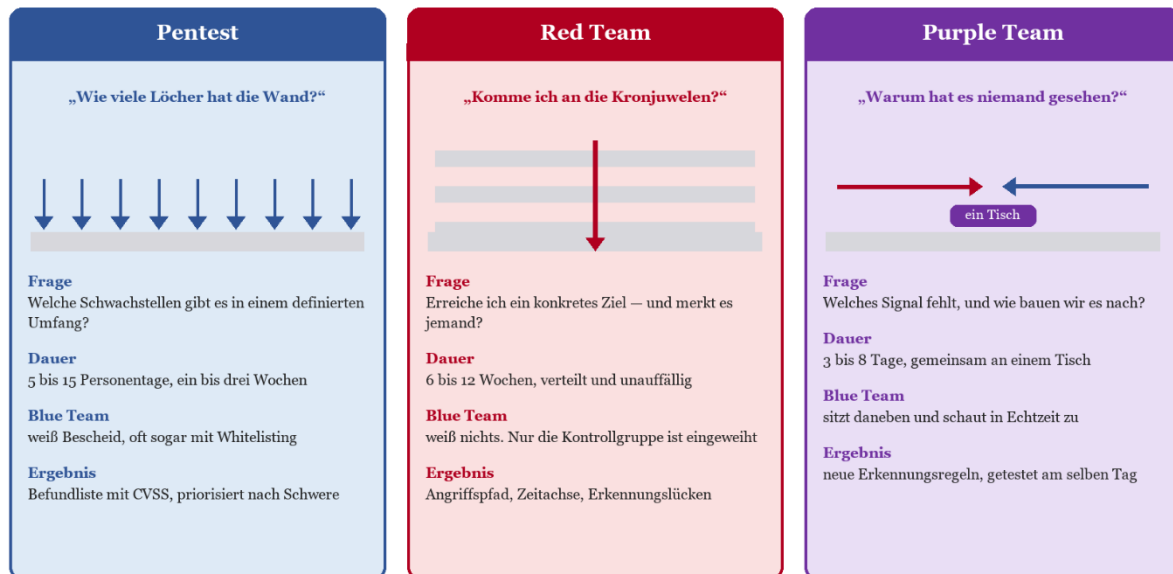
Die Microsoft Cloud Unified Penetration Testing Rules of Engagement verlangen seit Jahren keine Voranmeldung mehr. Sie verbieten aber unverändert: Denial-of-Service-Tests jeder Art, Zugriff auf Daten oder Tenants, die dir nicht gehören, die Verwendung fremder Zugangsdaten und Phishing gegen Beschäftigte von Microsoft. Findest du eine Schwachstelle in der Plattform selbst, gehört sie sofort an das Microsoft Security Response Center und nicht in deinen Abschlussbericht.

Für Finanzunternehmen ist die Übung seit DORA keine Kür mehr. Die Artikel 26 und 27 verlangen von den durch die Aufsicht benannten Instituten mindestens alle drei Jahre ein Threat-Led Penetration Testing auf Produktivsystemen. In Deutschland führt die BaFin die Aufsicht, die Deutsche Bundesbank begleitet operativ über TIBER-DE. Der komplette Zyklus aus Threat Intelligence, Red Teaming und Nachbesserung dauert sechs bis zwölf Monate.

Pentest, Red Team, Purple Team — drei Wörter, drei Rechnungen

Wenn drei Anbieter ein Angebot schicken und alle drei „Red Team“ darauf schreiben, heißt das nicht, dass du dreimal dasselbe bekommst. Es heißt in der Regel, dass zwei davon einen Pentest verkaufen und das Wort benutzen, weil es besser klingt. Das ist nicht durchweg Bosheit — die Begriffe sind tatsächlich unsauber im Markt. Aber du solltest sie sauber verwenden können, sonst kaufst du eine Befundliste und wunderst dich, warum niemand deine Erkennung getestet hat.

Drei Formate, drei Fragen — und drei sehr verschiedene Rechnungen



Skizze 1: Die drei Formate unterscheiden sich nicht im Werkzeugkasten, sondern in der Frage, die sie beantworten.

Der Pentest prüft die Wand

Ein Penetrationstest nimmt einen klar umrissenen Bereich — eine Webanwendung, ein Netzsegment, einen Tenant — und sucht darin systematisch nach Schwachstellen. Breite vor Tiefe. Das Ergebnis ist eine Liste mit Befunden, meist mit CVSS-Werten versehen, sortiert nach Schwere. Der Betrieb weiß in aller Regel Bescheid, oft ist die Test-IP sogar von der Filterung ausgenommen, damit der Tester in der verfügbaren Zeit wirklich prüfen kann und nicht die halbe Woche damit verbringt, an der Firewall vorbeizukommen.

Das ist keine Schwäche des Formats, sondern sein Zweck. Ein Pentest beantwortet die Frage: Wie viele Löcher hat diese Wand, und wie groß sind sie? Er beantwortet nicht die Frage, ob jemand merkt, dass gerade ein Loch benutzt wird. Wer einen Pentest kauft und Erkennungsqualität erwartet, hat das falsche Produkt gekauft.

Das Red Team prüft das Haus — und die Wachmannschaft

Ein Red Team bekommt ein Ziel, keinen Prüfumfang. „Zugriff auf das Postfach eines Vorstandsmitglieds“, „Auslesen der Rezepturdatenbank“, „Domänenadministrator in der Produktionsdomäne“. Und dann sucht es sich einen Weg. Es darf phishen, es darf einen Consent-Dialog missbrauchen, es darf ein Notebook im Homeoffice ins Visier nehmen, es darf geleakte

Zugangsdaten aus einem alten Datenleck ausprobieren. Es darf sogar warten — zwei Wochen nichts tun ist ein legitimer Teil des Vorgehens, wenn es dabei hilft, unter der Erkennungsschwelle zu bleiben.

Entscheidend ist: Das Blue Team weiß nichts. Nur die sogenannte Kontrollgruppe, drei bis fünf Personen, ist eingeweiht. Denn die eigentliche Frage lautet nicht „gibt es einen Weg?“ — den gibt es fast immer —, sondern „wie weit kommt jemand, bevor es auffällt?“ und „was passiert, wenn es auffällt?“. Ein Red Team misst nicht deine Schwachstellen. Es misst deine Erkennung, deine Reaktion und deine Organisation unter echten Bedingungen.

Das Purple Team ist der Teil, der wirklich etwas verändert

Purple Team ist kein drittes Team, sondern ein Arbeitsmodus: Angreifer und Verteidiger sitzen gemeinsam an einem Tisch. Das Red Team führt eine Technik aus und sagt laut, was es gerade tut. Das Blue Team schaut in seine Portale und sagt, was es sieht — oder eben nicht sieht. Fehlt das Signal, wird noch am selben Tag eine Erkennungsregel gebaut und die Technik wiederholt. Kein Bericht, keine Wartezeit, kein Ticketsystem dazwischen.

Wenn ich Häusern nur ein Format empfehlen dürfte, wäre es dieses. Ein Purple-Team-Tag bringt in der Regel mehr messbare Verbesserung pro ausgegebenem Euro als eine ganze Red-Team-Kampagne, einfach weil der Lernkanal direkt ist. Die Red-Team-Kampagne braucht man trotzdem: Sie ist der einzige Weg herauszufinden, ob die Regeln auch dann greifen, wenn niemand danebensteht und sagt, worauf zu achten ist. Die sinnvolle Reihenfolge lautet: härten, dann Purple, dann Red.

Kriterium	Pentest	Red Team	Purple Team
Leitfrage	Welche Schwachstellen gibt es hier?	Erreiche ich das Ziel, und merkt es jemand?	Welches Signal fehlt, und wie bauen wir es?
Umfang	eng definiert, technisch abgegrenzt	das ganze Haus, alle Wege	eine Liste von Techniken, systematisch
Dauer	5 bis 15 Personentage	6 bis 12 Wochen Laufzeit	3 bis 8 Tage
Blue Team	eingeweiht, oft mit Ausnahmen	ahnungslos, nur Kontrollgruppe weiß es	sitzt daneben und arbeitet mit
Erfolgsmaß	Anzahl und Schwere der Befunde	Erkennungsquote und Reaktionszeit	neue Regeln, sofort verifiziert
Ergebnis	Befundliste mit Bewertung	Angriffspfad mit Zeitachse	Erkennungsmatrix mit Abdeckungsgrad
Typische Kosten	8.000 bis 25.000 Euro	40.000 bis 120.000 Euro	10.000 bis 30.000 Euro
Sinnvoll, wenn	Grundhärtung steht, Nachweis nötig	Erkennung existiert und soll geprüft werden	Erkennung soll gezielt aufgebaut werden

Tabelle 1: Die drei Formate im Vergleich. Kosten sind Marktpannen für den deutschen Mittelstand, Stand September 2026.

WARNUNG — Der teuerste Fehler in der Reihenfolge

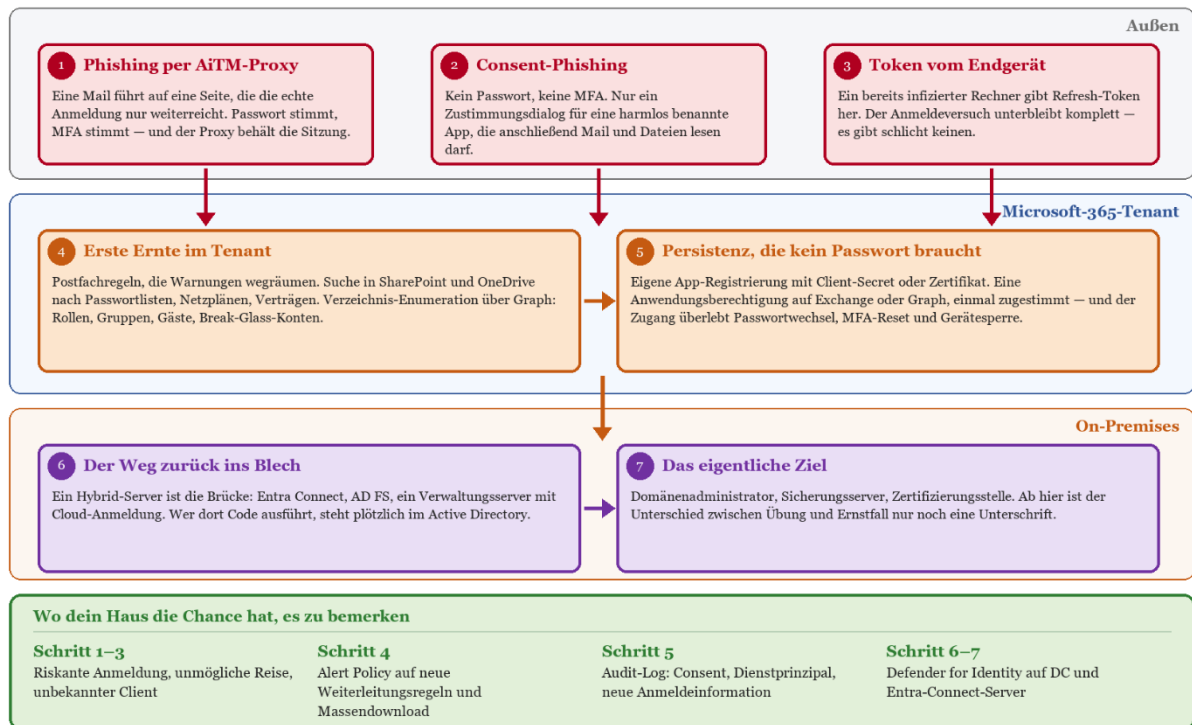
Ein Red Team gegen ein Haus laufen zu lassen, in dem Legacy Authentication noch offen ist, Consent für Endnutzer erlaubt bleibt und niemand die Audit-Logs liest, ist verschwendetes Geld. Das Ergebnis steht vorher fest, es kostet trotzdem fünfstellig und erzeugt einen Bericht, dessen erste zwanzig Seiten das wiederholen, was in jeder Härtungscheckliste steht.

Wenn du dir unsicher bist, wo dein Haus steht: Arbeite zuerst [Tenant-Härtung: die 20 wichtigsten Schrauben](#) durch und schalte [Legacy Authentication endgültig abschalten](#) ab. Ein Red Team lohnt sich ab dem Punkt, an dem du ehrlich sagen kannst: Wir glauben, wir würden das merken.

Was ein M365-fokussiertes Red Team tatsächlich anfasst

Der klassische Red-Team-Werkzeugkasten stammt aus der Zeit, in der Angriffe im Netz stattfanden: Perimeter überwinden, Fuß fassen, Rechte ausweiten, seitwärts bewegen. Das gibt es alles noch. Aber wenn deine Identitäten in Entra ID liegen, deine Dokumente in SharePoint und deine Postfächer in Exchange Online, verschiebt sich der Schwerpunkt. Der Perimeter ist die Anmeldung. Und die steht im Internet, für alle erreichbar, rund um die Uhr.

Der Pfad, den ein M365-Red-Team üblicherweise nimmt



Skizze 2: Der übliche Weg — drei mögliche Einstiege, eine gemeinsame Fortsetzung, und am Ende der Rückweg ins eigene Rechenzentrum.

Phishing, aber nicht das aus der Awareness-Schulung

Die Phishing-Mail, die ein Red Team schreibt, hat nichts mit dem generischen Paketdienst-Mist zu tun, den deine Belegschaft täglich sieht. Sie bezieht sich auf ein laufendes Projekt, das aus einer Stellenanzeige oder einem Presstext hervorgeht. Sie kommt von einer Domain, die deiner sehr ähnlich sieht und seit acht Wochen registriert ist, damit sie kein Neuling mehr ist. Sie geht an vier Personen, nicht an vierhundert, und zwar an die, deren Rolle die Mail plausibel macht.

Technisch dahinter steht meist ein Adversary-in-the-Middle-Proxy. Die Seite, auf der das Opfer landet, ist keine Fälschung, sondern ein Weiterleiter: Sie reicht Benutzernamen, Passwort und sogar die MFA-Bestätigung an die echte Microsoft-Anmeldung durch. Alles funktioniert, das Opfer landet tatsächlich in seinem Postfach — nur behält der Proxy das Sitzungs-Cookie. MFA hat nicht versagt, sie wurde umgangen. Deshalb ist die Kombination aus phishing-resistenten Verfahren und gerätegebundenen Tokens die einzige belastbare Antwort. Wie du das aufziehst, steht in [Passwordless: FIDO2 und Passkeys in der Praxis](#) und in [Conditional Access: Baselines, die nicht nerven](#).

Die Sitzungssteuerung Token Protection in Conditional Access — im Markt oft Token-Bindung genannt — bindet einen Token an das Gerät, für das er ausgestellt wurde. Ein geklautes Cookie auf

der Maschine des Angreifers ist damit wertlos. Für Windows ist die Funktion allgemein verfügbar, für macOS und iOS ist sie Anfang 2026 in die öffentliche Vorschau gegangen. Ergänzend sorgt Continuous Access Evaluation dafür, dass kritische Ereignisse wie eine Kontosperrung innerhalb von Minuten wirken statt erst beim nächsten Tokenwechsel; in der strikten Variante werden auch IP-basierte Standortrichtlinien laufend nachbewertet. Der Preis dafür ist eine Tokenlaufzeit von bis zu 28 Stunden statt der gewohnten 60 bis 90 Minuten — was viele Häuser beim ersten Blick ins Protokoll erschreckt und dann doch die sicherere Variante ist.

Consent: der Angriff, der gar keine Anmeldung braucht

Consent-Phishing ist die eleganteste Variante, weil sie das Passwort schlicht überspringt. Das Opfer klickt einen Link, sieht einen echten Microsoft-Zustimmungsdialog — nicht nachgebaut, sondern den originalen — und stimmt einer Anwendung namens „Mail Backup Service“ oder „Document Sync“ zu. Ab diesem Moment liest eine fremde Anwendung Postfach und Dateien, ohne dass jemals ein Kennwort das Haus verlassen hätte. Passwortwechsel hilft nicht. MFA hilft nicht. Ein neues Notebook hilft nicht.

Der Midnight-Blizzard-Fall zeigt die zweite Ausbaustufe: Eine vergessene Legacy-Anwendung, der irgendwann einmal eine hochprivilegierte Person zugestimmt hatte, wurde zum Sprungbrett. Der Angreifer registrierte darüber eigene Anwendungen und vergab sich die Exchange-Online-Anwendungsberechtigung für den Vollzugriff auf Postfächer. Wenn dein Red Team gut ist, wird es genau danach suchen: nach Dienstprinzipalen, die seit vier Jahren niemand angefasst hat, und nach Anwendungsberechtigungen, die jemand aus Bequemlichkeit erteilt hat. Wie du dort aufräumst, steht in [App-Registrierungen und Consent-Governance](#).

Token-Diebstahl vom Endgerät

Der dritte Einstieg macht gar keinen Umweg über den Menschen. Ein Rechner ist bereits kompromittiert — über eine manipulierte Software, ein Browser-Add-in, eine ClickFix-Aufforderung im Browser —, und der Angreifer greift dort die Refresh-Tokens ab. Aus Sicht von Entra ID passiert dabei zunächst überhaupt nichts Auffälliges: Es gibt keinen fehlgeschlagenen Anmeldeversuch, kein neues Land, keinen unbekannten Client. Es gibt nur einen bereits gültigen Zugang, der weiterbenutzt wird.

Deshalb ist der Endpunkt der einzige Ort, an dem dieser Schritt überhaupt sichtbar wird. Ein Red Team, das diesen Weg wählt, testet damit implizit die Qualität deines Endpunktschutzes — Sensorabdeckung, Manipulationsschutz, Reaktionszeit. Wenn du hier Lücken hast, findest du sie so. Der Einstieg dazu steht in [Defender for Endpoint: Onboarding ohne Drama](#).

Lateral Movement im Hybrid: der Weg zurück ins Blech

Der interessanteste Teil einer M365-Übung im deutschen Mittelstand ist selten die Cloud. Er ist die Brücke. Fast jedes Haus hier ist hybrid: Es gibt noch Domänencontroller, es gibt einen Fileserver, es gibt eine Maschinensteuerung, die nur mit einem Dienstkonto von 2011 spricht. Und es gibt mindestens einen Server, der beide Welten kennt — Entra Connect, AD FS, ein Verwaltungssystem, das sowohl an der Domäne hängt als auch mit einem Cloud-Konto administriert wird.

Dieser Server ist das Scharnier, und Angreifer wissen das. Wer dort Code ausführen kann, steht mit einem Schritt im Active Directory und arbeitet ab dann mit den bekannten Techniken weiter: Ticketangriffe, Delegierungsmissbrauch, Zertifikatsvorlagen mit zu weiten Rechten. Wenn dich die Ticketseite genauer interessiert, findest du die Grundlagen im [Kompetenzbereich Kerberos](#).

Auf der Verteidigungsseite ist genau dafür Defender for Identity gedacht. Der Sensor der Version 3.x unterstützt inzwischen auch Domänencontroller, auf denen zusätzlich AD FS, AD CS oder Entra

Connect laufen — und Microsoft rollt Erkennungen und Empfehlungen für diese Rollen schrittweise aus, beginnend mit Entra Connect. Technisch nutzt der neue Sensor die Windows Filtering Platform für tiefere Netzwerkeinsicht und erweitertes RPC-Auditing, bei gedeckelten 30 Prozent CPU und 1,5 Gigabyte Arbeitsspeicher. Die Erkennung für einen möglichen Golden-Ticket-Angriff über verdächtige Tickets ist allgemein verfügbar. Wie du das sauber ausrollst, steht in [Defender for Identity: das Active Directory im Blick](#).

Technik im Red Team	Was dabei wirklich geprüft wird	Wo das Signal auftauchen müsste
Zielgerichtetes Phishing mit AiTM-Proxy	Melden Menschen verdächtige Mails? Greifen Safe Links? Fällt die neue Domain auf?	Defender for Office 365, Meldebutton, Postfach des Sicherheitsteams
Consent für eine fremde Anwendung	Darf die Belegschaft überhaupt zustimmen? Wird ein Antrag geprüft? Fällt der Dienstprinzipal auf?	Einheitliches Überwachungsprotokoll, Consent-Anfragen, App-Governance
Diebstahl von Refresh-Tokens vom Client	Sieht der Endpunktsensor die Ablage? Ist Manipulationsschutz aktiv? Greift Token-Bindung?	Defender for Endpoint, Anmeldeprotokoll ohne passenden Anmeldevorgang
Anlegen einer eigenen App-Registrierung	Wer darf Anwendungen registrieren? Wird eine neue Anmeldeinformation bemerkt?	Verzeichnisprotokoll, Alert Policy auf Anwendungsberechtigungen
Postfachregel, die Warnungen wegräumt	Existiert eine Regel dafür? Landet die Meldung bei jemandem, der reagiert?	Alert Policy im Defender-Portal, Weiterleitungsberichte
Massendownload aus SharePoint	Ist eine Schwelle definiert? Wird ungewöhnliches Verhalten überhaupt gemessen?	Defender for Cloud Apps, Aktivitätsrichtlinien
Zugriff auf den Entra-Connect-Server	Ist der Server wie ein Domänencontroller geschützt? Läuft dort ein Sensor?	Defender for Identity, privilegierte Verwaltungswege
Ticket- und Delegierungsangriffe im AD	Erkennt jemand ungewöhnliche Ticketanforderungen? Sind Delegierungen aufgeräumt?	Defender for Identity, Ereignisprotokolle der Domänencontroller

Tabelle 2: Jede Technik im Bericht ist eigentlich eine Frage an deine eigene Erkennung. Der Angriffspfad ist nur die Verpackung.

TIPP — Die Übung, die zehnmal weniger kostet und trotzdem viel zeigt

Bevor du ein Red Team beauftragst, spiele den Phishing-Teil selbst durch. Attack Simulation Training in Defender for Office 365 liefert dir echte Klickquoten, echte Meldequoten und die Erkenntnis, ob dein Meldebutton überhaupt irgendwo ankommt. Das Vorgehen steht in [Attack Simulation Training: Phishing-Übungen richtig aufsetzen](#).

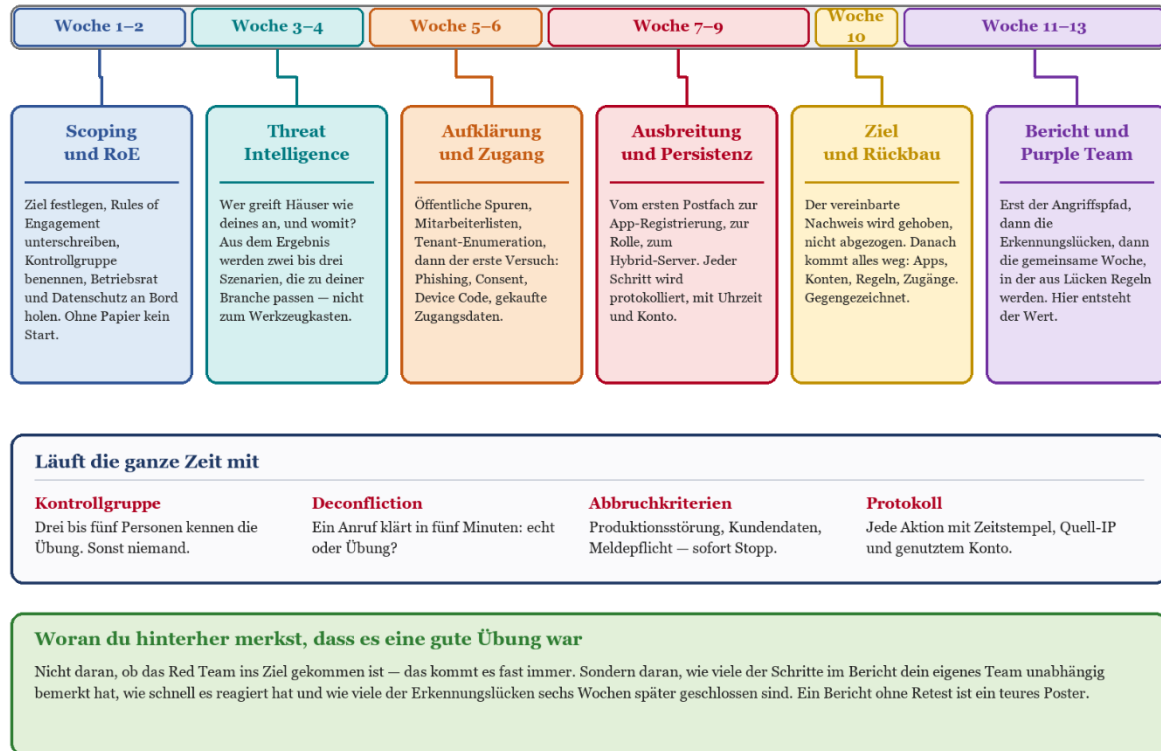
Und die organisatorische Seite übst du am grünen Tisch: eine [Feuerwehrübung: Security-Notfälle im Team trainieren](#) kostet 90 Minuten und einen Besprechungsraum. Wer diese beiden Vorübungen gemacht hat, holt aus einer Red-Team-Kampagne deutlich mehr heraus — weil dann die peinlichen Befunde schon weg sind und die interessanten übrig bleiben.

Scope und Rules of Engagement — der Vertrag, der die Übung rettet

Die Rules of Engagement sind das Dokument, an das sich beide Seiten erinnern, wenn etwas schiefgeht. Und irgendwann geht etwas schief: Ein Test läuft in ein System, das niemand auf dem Zettel hatte. Ein Dienstkonto wird gesperrt und ein Fertigungsauftrag bleibt stehen. Ein Mitarbeiter meldet die Phishing-Mail nicht intern, sondern bei der Polizei. Das sind keine

Katastrophen, wenn vorher aufgeschrieben wurde, wer dann was tut. Ohne dieses Papier sind es sehr wohl welche.

Ablauf einer Red-Team-Übung — dreizehn Wochen am Stück



Skizze 3: Der typische Ablauf einer Red-Team-Übung über dreizehn Wochen. Die letzten drei Wochen sind die, in denen der Wert entsteht.

Was in den Scope gehört — und was ausdrücklich nicht

Der Scope einer Red-Team-Übung ist nicht die Liste der erlaubten IP-Adressen, sondern die Beschreibung des Ziels und der Ausschlüsse. Formuliere das Ziel so konkret, dass am Ende unstrittig ist, ob es erreicht wurde. „Zugriff auf Postfach und OneDrive der Geschäftsführung“ ist ein Ziel. „Prüfung der Sicherheit von Microsoft 365“ ist eine Absichtserklärung.

Auf der Ausschlussseite gehören Systeme hinein, bei denen ein Ausfall Menschen gefährdet oder Produktion stoppt: Medizintechnik, Leitsysteme, alles mit Sicherheitsfunktion. Ebenso Personen, die aus guten Gründen nicht Ziel einer Simulation sein sollen — die Betriebsärztin, der Betriebsrat als Gremium, Konten in Elternzeit. Und schließlich physische Methoden: Wer nicht will, dass jemand nachts über den Zaun steigt oder sich als Handwerker ausgibt, muss das hinschreiben. Sonst ist es erlaubt, denn ein echter Angreifer würde es auch tun.

Die Spielregeln der Plattform

Für alles, was in der Microsoft-Cloud stattfindet, gilt zusätzlich das Regelwerk des Anbieters. Eine Voranmeldung braucht es nicht mehr — das war einmal anders und hält sich hartnäckig in Checklisten. Die Verbote gelten aber unverändert. Denial-of-Service ist unter allen Umständen untersagt, auch als Lasttest getarnt. Zugriff auf Tenants oder Speicherkonten, die dir nicht gehören, ist untersagt. Fremde Zugangsdaten zu benutzen, ist untersagt. Phishing gegen

Beschäftigte von Microsoft ist untersagt. Ausdrücklich erlaubt ist dagegen, deine eigene Erkennung und Überwachung zu testen — genau das, was du vorhast.

Ein zweiter Punkt wird gern übersehen: Findet dein Dienstleister eine Schwachstelle in der Plattform selbst und nicht in deiner Konfiguration, gehört sie unverzüglich an das Microsoft Security Response Center gemeldet. Nimm das in den Vertrag auf, samt der Frage, wem die Erkenntnis gehört. Und wenn versehentlich auf Daten zugegriffen wurde, die nicht deine sind: sofort aufhören, dokumentieren, melden.

Abbruchkriterien, Deconfliction und der Freibrief

Drei Mechanismen entscheiden darüber, ob eine Übung im Ernstfall beherrschbar bleibt. Erstens die Abbruchkriterien: Eine spürbare Produktionsstörung, ein Zugriff auf Kunden- oder Gesundheitsdaten, ein tatsächlicher Fremdangriff parallel zur Übung, das Auslösen einer aufsichtsrechtlichen Meldepflicht. In diesen Fällen wird nicht diskutiert, sondern gestoppt.

Zweitens die Deconfliction: eine Nummer, unter der die Kontrollgruppe rund um die Uhr klären kann, ob eine beobachtete Aktivität zur Übung gehört oder echt ist. Diese Frage kommt, und sie kommt selten zu einer angenehmen Uhrzeit. Wenn dein Sicherheitsteam mitten in der Nacht einen Vorfall ausruft und die halbe Firma weckt, weil das Red Team gerade arbeitet, hast du kein Erkenntnisproblem, sondern ein Prozessproblem — allerdings eines, das den Ablauf deines [Incident-Response-Plan für Microsoft 365](#) sehr präzise offenlegt.

Drittens der Freibrief. Ein einseitiges Blatt mit Firmenlogo, Auftragsnummer, Zeitraum, Namen der Testenden und zwei Telefonnummern, das jede Person im Team körperlich bei sich trägt. Das ist kein Ritual. Ohne unterschriebenen Auftrag ist das, was ein Red Team tut, das Ausspähen von Daten nach Paragraf 202a und je nach Wirkung Computersabotage nach Paragraf 303b des Strafgesetzbuchs. Der Auftrag ist die Rechtfertigung, und die muss vorzeigbar sein.

Punkt in den Rules of Engagement	Konkret hineinschreiben	Warum
Ziel der Übung	ein bis zwei Sätze, überprüfbar formuliert	sonst streitet ihr am Ende über den Erfolg
Zeitfenster	Kalenderzeitraum, erlaubte Tageszeiten, Sperrtermine	Jahresabschluss und Inventur sind schlechte Übungswochen
Ausgeschlossene Systeme	namentlich, mit Kennung	„kritische Systeme“ ist keine Abgrenzung
Ausgeschlossene Personen	Rollen und Einzelfälle	Fürsorgepflicht und Betriebsratsvereinbarung
Erlaubte Methoden	Phishing ja, physisch nein, Anrufe nur ohne Druck	was nicht ausgeschlossen ist, gilt als erlaubt
Abbruchkriterien	vier bis sechs Fälle, abschließend	im Ernstfall wird nicht neu verhandelt
Deconfliction	zwei Namen, zwei Nummern, rund um die Uhr	die Frage kommt nachts um halb drei
Protokollpflicht	Zeitstempel, Quell-IP, genutztes Konto, Aktion	sonst ist die spätere Forensik unmöglich
Rückbau	Liste aller Artefakte, gegengezeichnet	Hintertüren aus einer Übung sind echte Hintertüren
Datenumgang	Nachweis statt Kopie, Löschfrist, Verschlüsselung	Datenschutz und Berufsgeheimnisse
Plattformregeln	Verweis auf die Microsoft-Regeln, Meldeweg an MSRC	Verstöße können den Tenant kosten
Berichtsform	Managementteil, technischer Teil, Erkennungsmatrix	sonst kommt eine Rohdatensammlung

Tabelle 3: Die Mindestpunkte für die Rules of Engagement. Fehlt einer davon, verhandelst du ihn später unter Zeitdruck.

WICHTIG — Der Rückbau ist Teil der Leistung, nicht eine nette Geste

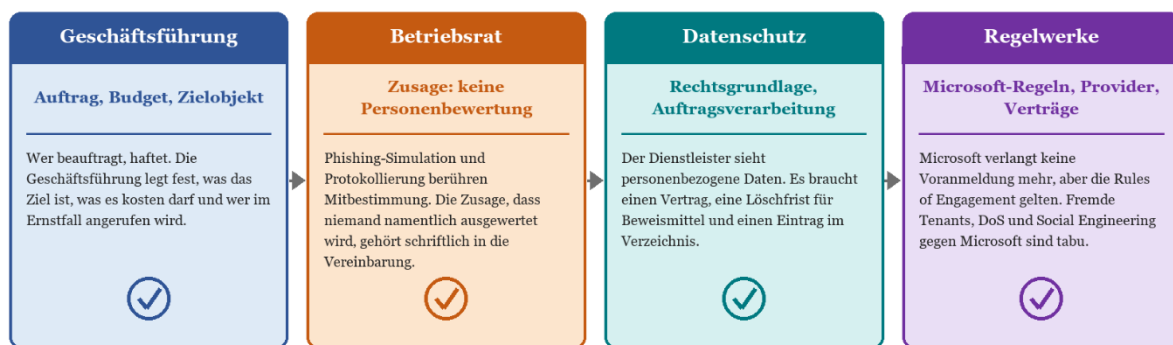
Jede App-Registrierung, jedes Client-Secret, jede Postfachregel, jedes angelegte Konto, jeder hinterlegte Schlüssel und jedes ausgerollte Werkzeug muss am Ende weg sein — und zwar nachweislich, mit einer Liste, die beide Seiten unterschreiben.

Ich habe mehr als einmal Umgebungen gesehen, in denen ein Dienstprinzipal aus einem Test von vor drei Jahren noch mit Anwendungsberechtigungen im Verzeichnis stand. Der Unterschied zwischen einer Übungshintertür und einer echten Hintertür ist ausschließlich die Absicht dessen, der den Schlüssel gerade in der Hand hält.

Betriebsrat und Datenschutz — vorher, nicht nachher

Hier scheitern in Deutschland mehr Red-Team-Projekte als an der Technik. Nicht, weil Betriebsrat oder Datenschutzbeauftragte etwas gegen Sicherheit hätten — das Gegenteil ist die Regel —, sondern weil sie zu spät gefragt werden. Wer drei Wochen vor Projektstart mit einem fertigen Vertrag ankommt und um eine schnelle Zustimmung bittet, bekommt zu Recht eine Rückfrage und im schlimmsten Fall eine Blockade. Wer sechs Monate vorher mit einer leeren Seite kommt und fragt, wie man das gemeinsam sauber hinbekommt, bekommt einen Verbündeten.

Vier Freigaben, bevor die erste Mail rausgeht



Fehlt eine davon, ist es keine Übung

Ein Red Team ohne unterschriebenen Auftrag begeht Computersabotage nach Paragraph 303b und Ausspähen von Daten nach Paragraph 202a. Der berühmte „Freibrief“ — ein einseitiges Blatt mit Unterschrift und Telefonnummer, das jeder Tester bei sich trägt — ist kein Ritual, sondern die einzige Antwort auf einen Werkschutz um drei Uhr nachts.

Skizze 4: Vier Freigaben. Wer eine davon überspringt, hat keine Übung durchgeführt, sondern ein Risiko.

Warum das Thema die Mitbestimmung berührt

Der Kern ist Paragraph 87 Absatz 1 Nummer 6 des Betriebsverfassungsgesetzes: Einrichtungen, die geeignet sind, Verhalten oder Leistung der Belegschaft zu überwachen, sind mitbestimmungspflichtig — unabhängig davon, ob du sie so einsetzen willst. Eine Phishing-Simulation misst, wer geklickt hat. Ein Red-Team-Protokoll hält fest, über welches Konto der Einstieg gelang. Die Auswertung der Anmeldeprotokolle während der Übung zeigt Arbeitszeiten. Alles drei ist geeignet.

Die Lösung ist selten ein Verbot, sondern fast immer eine Zusage: Die Auswertung erfolgt ausschließlich aggregiert. Namen werden im Bericht pseudonymisiert. Aus einem Klick folgt keine arbeitsrechtliche Konsequenz, sondern höchstens ein Schulungsangebot. Die Rohdaten werden nach einer festen Frist gelöscht. Wenn diese vier Sätze in der Betriebsvereinbarung stehen, ist das

Gespräch in der Regel kurz. Der ausführliche Rahmen dazu steht in [Security-Monitoring und Betriebsrat: was mitbestimmungspflichtig ist](#).

WICHTIG — Die vier Sätze, die das Gespräch mit dem Betriebsrat entspannen

Erstens: Es wird keine personenbezogene Erfolgsquote erhoben oder berichtet. Auswertungen erfolgen ausschließlich auf Ebene von Abteilungen ab einer Mindestgröße.

Zweitens: Ein Klick auf eine simulierte Phishing-Mail führt zu keiner Abmahnung, keinem Eintrag in die Personalakte und keinem Gespräch mit der Führungskraft.

Drittens: Die während der Übung erhobenen Protokolle werden spätestens 90 Tage nach Abnahme des Berichts gelöscht; die Löschung wird bestätigt.

Viertens: Der Betriebsrat erhält denselben Abschlussbericht wie die Geschäftsführung und wird auf Wunsch vorab über den Zeitraum informiert.

Was die Datenschutzseite braucht

Der Dienstleister wird im Zuge der Übung personenbezogene Daten sehen: Namen, Mailadressen, Postfachinhalte, unter Umständen Gesundheits- oder Beschäftigtendaten. Das ist Auftragsverarbeitung nach Artikel 28 der Datenschutz-Grundverordnung und braucht einen entsprechenden Vertrag mit den üblichen Anlagen: technische und organisatorische Maßnahmen, Unterauftragnehmer, Löschkonzept.

Die Rechtsgrundlage für die Übung selbst ist regelmäßig das berechtigte Interesse nach Artikel 6 Absatz 1 Buchstabe f — die Sicherheit der Verarbeitung ist in Artikel 32 sogar ausdrücklich gefordert. Dokumentiere die Abwägung schriftlich, statt sie nur zu denken. Nimm die Verarbeitung ins Verzeichnis auf. Und kläre vorab, ob eine Datenschutz-Folgenabschätzung nötig ist: Bei einer Übung mit systematischer Auswertung von Beschäftigtenverhalten ist die Frage ernsthaft zu stellen und die Antwort zu begründen — in beide Richtungen.

Der wichtigste operative Punkt heißt Nachweis statt Kopie. Wenn das Ziel „Zugriff auf die Personalakten“ lautet, ist der Nachweis ein Bildschirmfoto mit geschwärzten Inhalten und einer Datei-Kennung — nicht ein Abzug des Ordners auf das Notebook des Testers. Wer Berufsgeheimnisse verwaltet, Kanzlei, Praxis, Steuerberatung, sollte diese Regel wörtlich in den Vertrag schreiben; Paragraph 203 des Strafgesetzbuchs kennt keinen Übungsvorbehalt. Wie du Fundstellen und Klassifizierung sauber führst, gehört in den [Kompetenzbereich Microsoft Purview](#).

WARNUNG — Der Moment, in dem aus der Übung ein meldepflichtiger Vorfall wird

Wenn während der Übung ein echter Angreifer auffällt — was häufiger vorkommt, als man denkt, weil plötzlich jemand genau hinschaut —, endet die Übung sofort und der Vorfallsprozess beginnt. Das ist ein Erfolg, kein Betriebsunfall.

Und wenn die Übung selbst eine Störung verursacht, die Dienste erheblich beeinträchtigt, kann für betroffene Einrichtungen die NIS2-Meldekette greifen: Frühwarnung binnen 24 Stunden, Folgemeldung binnen 72 Stunden, Abschlussbericht binnen eines Monats. Wer davon betroffen ist und was das technisch bedeutet, steht in [NIS2 aus Security-Sicht: technische Maßnahmen in M365](#). Genau deshalb steht dieser Fall in den Abbruchkriterien.

Kosten, Aufwand und was du aus dem Bericht machst

Reden wir über Geld, denn an dieser Stelle hört das Gespräch in vielen Häusern auf. Die Zahlen sind unangenehm, aber sie sind nicht geheim, und sie sind auch nicht willkürlich. Tagessätze zertifizierter Testerinnen und Tester liegen in Deutschland im September 2026 üblicherweise

zwischen 1.200 und 2.000 Euro; über den breiten Markt gerechnet beginnt die Spanne bei etwa 900 Euro. Ein Web-Anwendungstest mit acht Testtagen und zwei Tagen Berichterstellung landet damit bei rund 13.000 Euro. Ein vollständiges Red-Team-Assessment beginnt bei etwa 30.000 Euro und reicht bis 80.000 Euro und darüber, wenn Threat Intelligence, physische Komponenten und eine lange Laufzeit dazukommen.

Was diese Zahlen verschweigen, ist der interne Aufwand. Der ist erheblich und wird fast nie eingeplant. Rechne bei einer mittelgroßen Kampagne mit zehn bis fünfzehn internen Personentagen: Scoping, Vertragsabstimmung, Betriebsrat, Datenschutz, Begleitung der Kontrollgruppe, Berichtsbesprechung. Die Kontrollgruppe muss über die gesamte Laufzeit erreichbar sein — nicht voll ausgelastet, aber ansprechbar. Und die Nacharbeit ist der größte Posten von allen: Ein Bericht mit dreißig Befunden erzeugt schnell dreißig bis sechzig Personentage Arbeit, wenn du ihn ernst nimmst.

Format	Externe Kosten	Laufzeit	Interner Aufwand	Wofür es taugt
Phishing-Simulation intern	Lizenzkosten, sonst nichts	2 bis 4 Wochen	3 bis 5 Personentage	Klickquote, Meldequote, erste Ehrlichkeit
Tabletop-Übung	0 bis 5.000 Euro	ein halber Tag	2 bis 4 Personentage	Entscheidungswege, Rollen, Erreichbarkeit
Konfigurationsprüfung Tenant	5.000 bis 12.000 Euro	1 bis 2 Wochen	2 bis 4 Personentage	Härtungslücken vor der teuren Übung
Pentest M365 und Perimeter	8.000 bis 25.000 Euro	2 bis 4 Wochen	4 bis 8 Personentage	Nachweis für Kunden, Auditoren, Versicherung
Purple-Team-Woche	10.000 bis 30.000 Euro	1 bis 2 Wochen	5 bis 10 Personentage	gezielter Aufbau von Erkennung
Red-Team-Kampagne	30.000 bis 80.000 Euro	8 bis 14 Wochen	10 bis 15 Personentage	Realitätsprüfung für Erkennung und Reaktion
TLPT nach DORA oder TIBER-DE	100.000 Euro aufwärts	6 bis 12 Monate	25 Personentage aufwärts	aufsichtsrechtliche Pflicht im Finanzsektor

Tabelle 4: Marktspannen für den deutschen Mittelstand, Stand September 2026. Der interne Aufwand ist ohne die Umsetzung der Maßnahmen gerechnet — die kommt obendrauf.

Der Bericht ist kein Ergebnis, sondern Rohmaterial

Ein guter Red-Team-Bericht hat drei Teile. Der erste ist eine Erzählung: Wie kam der Angriff herein, welchen Weg nahm er, wann wurde welches Ziel erreicht — mit Zeitstempeln, damit du die eigene Erkennung dagegenlegen kannst. Der zweite Teil ist die Erkennungsmatrix: welche Technik wurde ausgeführt, gab es ein Signal, wurde es bemerkt, wurde reagiert. Der dritte Teil ist die Befundliste mit Empfehlungen. Fehlt der zweite Teil, hast du einen teuren Pentest gekauft.

Was aus 34 Befunden am Ende wirklich wird



Skizze 5: Aus dreißig Seiten Bericht werden vier Maßnahmen mit Namen und Datum. Alles andere ist Ablage.

Die Arbeit beginnt danach. Entdoppeln zuerst: Drei Befunde mit derselben Ursache sind ein Befund mit drei Fundorten. Dann bewerten — und zwar nicht nach der Punktzahl aus dem Werkzeug, sondern nach der Frage, wie viel Aufwand ein Angreifer für die Ausnutzung braucht und was er dann erreicht. Ein mittelschwerer Befund, der ohne Vorbedingung Domänenadministrator liefert, schlägt drei kritische, die eine bereits kompromittierte Maschine voraussetzen.

Am Ende bleiben typischerweise drei bis fünf Maßnahmen übrig, die sofort passieren, jede mit einer verantwortlichen Person und einem Datum. Der Rest wandert in ein Backlog, das im Quartalsrhythmus abgearbeitet wird. Und dann kommt der Schritt, den fast alle auslassen: der Retest nach sechs bis acht Wochen. Zwei bis drei Tage, in denen genau die geschlossenen Wege noch einmal gegangen werden. Erst danach darf im [Security-Reporting für die Geschäftsführung](#) an die Geschäftsführung „erledigt“ stehen.

Die Nacharbeit gehört ins Purple Team

Der wertvollste Teil des Berichts ist die Spalte, in der steht: keine Erkennung. Jede Zeile darin ist ein Auftrag. Setz dich mit dem Red Team für drei Tage zusammen, lass jede unerkannte Technik erneut ausführen und baue währenddessen die Regel. Bei Microsoft-Werkzeugen heißt das konkret: eine Alert Policy nachschärfen, eine Aktivitätsrichtlinie in Defender for Cloud Apps anlegen, eine Abfrage in der erweiterten Suche schreiben und als benutzerdefinierte Erkennungsregel planen. Der Einstieg dazu steht in [Alert Policies und Benachrichtigungen richtig aufsetzen](#) und in [Defender for Cloud Apps im Mittelstand](#). Wenn du die Signale zusammenführen und länger aufbewahren willst, ist der Weg über [Microsoft Sentinel anbinden: wann es sich lohnt](#) beschrieben.

Und noch ein Punkt, der selten im Angebot steht: Lass das Red Team eine Stunde vor deinem Sicherheitsteam sprechen und den Angriff selbst erzählen. Ohne Folien, ohne Vorwurf, mit Bildschirmaufnahmen. Ich habe in diesen Stunden mehr Aha-Momente erlebt als in jedem Bericht. Wenn dein Team danach auch noch die Verteidigungsseite systematisch nacharbeiten soll, ist eine [Schulung Microsoft 365 Security](#) der schnellere Weg als ein weiteres Werkzeug.

WEITER — Passend zu diesem Beitrag

- › Vorher üben, ohne dass jemand angreift: [Feuerwehrübung: Security-Notfälle im Team trainieren](#)
- › Die Sensorik, die den Hybrid-Teil überhaupt sichtbar macht: [Defender for Identity: das Active Directory im Blick](#)
- › Der rechtliche Rahmen für Protokollierung und Auswertung: [Security-Monitoring und Betriebsrat: was mitbestimmungspflichtig ist](#)
- › Der Plan, den die Übung auf die Probe stellt: [Incident-Response-Plan für Microsoft 365](#)
- › Aufräumen bei Dienstprinzipalen und Zustimmungen: [App-Registrierungen und Consent-Governance](#)
- › Was passiert, wenn es kein Test war: [Dossier TerminalFix und der Berlin-Hack](#)
- › Begleitung bei Scope, Ausschreibung und Nacharbeit: [Microsoft 365 Security Beratung](#)

Häufige Fragen

Wir sind 150 Leute. Ist ein Red Team für uns nicht völlig überzogen?

In den meisten Fällen ja — jedenfalls als erster Schritt. Bei 150 Beschäftigten ist die sinnvolle Reihenfolge: Tenant härten, Legacy-Authentifizierung abschalten, Conditional Access ordentlich aufsetzen, eine Phishing-Simulation fahren, eine Tabletop-Übung machen. Das kostet zusammen einen Bruchteil und findet neunzig Prozent dessen, was ein Red Team auch finden würde. Interessant wird eine echte Kampagne ab dem Punkt, an dem du eine funktionierende Erkennung hast und wissen willst, ob sie unter realen Bedingungen greift. Vorher misst du nur, dass ungeschützte Türen offen sind.

Muss ich den Test bei Microsoft anmelden?

Nein. Die Voranmeldung ist seit Jahren abgeschafft, die Regel hält sich aber hartnäckig in älteren Checklisten. Was bleibt, sind die inhaltlichen Vorgaben der Rules of Engagement: keine Denial-of-Service-Tests, kein Zugriff auf fremde Tenants oder Daten, keine fremden Zugangsdaten, kein Social Engineering gegen Beschäftigte von Microsoft. Ausdrücklich erlaubt ist, deine eigene Erkennung und Überwachung zu testen. Nimm eine Zusicherung der Regeltreue in den Vertrag mit deinem Dienstleister auf — Verstöße können im Extremfall den Tenant kosten, und die Rechnung trägst dann du, nicht der Dienstleister.

Wie viele Leute im Haus dürfen von der Übung wissen?

So wenige wie möglich, aber die richtigen. Bewährt haben sich drei bis fünf: die Geschäftsführung oder deren Vertretung, die IT-Leitung, eine Person aus der Sicherheit, die nicht selbst im operativen Erkennungsbetrieb sitzt, und je nach Haus die Person aus dem Datenschutz. Der Betriebsrat wird über den Rahmen informiert, nicht über den konkreten Zeitpunkt — das lässt sich in der Vereinbarung so festhalten. Alle anderen erfahren es hinterher. Jede zusätzliche eingeweihte Person senkt den Erkenntniswert, und zwar messbar: Menschen verhalten sich anders, wenn sie wissen, dass geschaut wird.

Was, wenn das Red Team gar nicht reinkommt?

Dann hast du entweder ein wirklich gutes Haus oder ein zu enges Zeitfenster. Ehrlicherweise ist Letzteres häufiger. Für diesen Fall gehört ins Angebot eine Eskalationsstufe, die im Fachjargon Assumed Breach heißt: Wenn nach der Hälfte der Zeit kein Einstieg gelungen ist, bekommt das Team ein normales Benutzerkonto und einen verwalteten Client gestellt und arbeitet von dort weiter. Das ist kein Trostpreis, sondern realistisch — irgendwann klickt jemand, und die

interessante Frage ist ohnehin, was danach passiert. Vereinbare diesen Übergang vorher, sonst verhandelst du ihn mitten im Projekt.

Wie oft sollten wir das wiederholen?

Für die meisten Häuser außerhalb regulierter Branchen: eine große Kampagne alle zwei bis drei Jahre, dazwischen jährlich ein kleineres Format — Purple-Team-Tage, eine Konfigurationsprüfung, eine Phishing-Simulation. Wichtiger als die Frequenz ist der Anlass: Nach einer Migration, nach einer Zukäufen mit fremdem Tenant, nach einem Wechsel im Sicherheitsteam oder nach der Einführung eines neuen Erkennungswerkzeugs ist eine Überprüfung sinnvoller als nach Kalender. Im Finanzsektor gibt DORA den Takt vor: mindestens alle drei Jahre für die benannten Institute.

Unser Dienstleister will Domänenadministrator, um Zeit zu sparen. Sinnvoll?

Für einen Konfigurationsaudit ja, für ein Red Team nein. Wer mit vollen Rechten startet, prüft nichts mehr am Weg dorthin — und der Weg dorthin ist genau das Produkt. Ein sinnvoller Mittelweg: Die Kampagne beginnt ohne Rechte, und parallel läuft ein getrennt beauftragter Konfigurationsaudit mit lesendem Zugriff. Beides zusammen ergibt ein vollständiges Bild, ohne dass eines das andere entwertet. Nur bitte nicht dasselbe Team zeitgleich mit beidem beauftragen, sonst weiß hinterher niemand mehr, was tatsächlich gefunden und was einfach abgelesen wurde.

Wie verhindere ich, dass der Bericht Menschen an den Pranger stellt?

Indem du es vorher schriftlich ausschließt und den Bericht vor der Verteilung liest. Konkret: keine Namen im Bericht, sondern Rollen und Pseudonyme. Keine Aufzählung, wer geklickt hat, sondern Angaben auf Abteilungsebene ab einer Mindestgröße. Und im Vorwort ein Satz der Geschäftsführung, dass die Übung Prozesse und Technik prüft, nicht Personen. Das ist nicht Kuschelpädagogik, sondern Eigennutz: In einem Haus, in dem ein Klick Ärger bedeutet, meldet niemand mehr eine verdächtige Mail. Und die Meldequote ist die einzige Kennzahl aus dem Awareness-Bereich, die im Ernstfall wirklich hilft.

Was kostet uns die Übung an Ausfallrisiko?

Bei sauberer Vorbereitung wenig, aber nicht null. Die realistischen Störungen sind: ein gesperrtes Dienstkonto nach einem Password Spray, eine überlastete Anmeldung bei zu schnellem Vorgehen, ein ausgelöster Vorfallsprozess mit entsprechender Aufregung. Alle drei lassen sich durch Vorgaben in den Rules of Engagement stark eindämmen — langsames Spraying mit definierter Rate, Ausschluss von Dienstkonten, Deconfliction-Nummer. Was du nicht ausschließen kannst und auch nicht solltest: dass dein Team nachts arbeitet, weil es glaubt, es sei ernst. Genau dafür ist die Übung da.

Können wir das nicht selbst machen?

Den Purple-Team-Teil zunehmend ja, den Red-Team-Teil praktisch nie. Nicht, weil die Fähigkeiten fehlen müssten, sondern weil ein internes Team seine eigene Umgebung kennt und damit die wichtigste Eigenschaft eines Angreifers verliert: Unwissenheit. Dazu kommt das saubere Vier-Augen-Prinzip — wer die Erkennung gebaut hat, sollte sie nicht selbst prüfen. Was du hervorragend selbst kannst: Phishing-Simulationen, Konfigurationsprüfungen anhand des Secure Score, Tabletop-Übungen und die Nacharbeit. Das ist der überwiegende Teil des Nutzens, zu einem kleinen Teil der Kosten.

Fazit

Eine Red-Team-Übung ist kein Sicherheitsprodukt, sondern eine Messung. Sie sagt dir nicht, ob du sicher bist — diese Frage hat keine Antwort —, sondern wie weit jemand kommt, bevor du es merkst, und was passiert, wenn du es merkst. Das ist die einzige Zahl in der IT-Sicherheit, die sich nicht schönrechnen lässt, und deshalb ist sie so unbequem und so wertvoll.

Der Rest ist Handwerk. Die Begriffe sauber trennen, damit du das kaufst, was du brauchst: Pentest für die Wand, Purple Team für die Erkennung, Red Team für die Realitätsprüfung. Die Reihenfolge einhalten — härten, dann üben, dann angreifen lassen. Das Ziel so formulieren, dass am Ende niemand über den Erfolg streiten muss. Die Rules of Engagement ernst nehmen, einschließlich der Abbruchkriterien und des Rückbaus. Betriebsrat und Datenschutz früh einbinden, weil sie sonst zu Recht bremsen. Und den Bericht als Anfang behandeln, nicht als Ergebnis.

Wenn du am Ende genau vier Maßnahmen umsetzt, jede mit einem Namen und einem Datum, und nach acht Wochen einen Retest machst, hat sich die Übung gelohnt. Wenn der Bericht in einem Ordner landet und dort auf das nächste Audit wartet, hast du fünfstellig für ein Dokument bezahlt, das im Ernstfall niemand aufschlägt. Der Unterschied zwischen beiden Fällen liegt nicht im Dienstleister, nicht im Werkzeug und nicht im Budget. Er liegt in den vier Wochen nach der Abnahme.

Und noch etwas: Wenn dein Team während der Übung mitten in der Nacht die halbe Firma weckt, weil es glaubt, es sei ernst — dann feiere das. Genau dieses Verhalten willst du. Der Rest ist Feinschliff, und Feinschliff kann man üben.