Heft 38 | Dezember 2023 | 17. Jahrgang | ISSN 1864-6557

IT-GOVERNANCE

Fachzeitschrift des ISACA Germany Chapter e.V.

Zum Thema:

Nach Microsoft-Schlüssel-Diebstahl aus Crash-Dump – neues IT-Audit-Prüfthema: IT-Diagnose-Daten

Stephen Fedtke

Sonderdruck

ISACA Germany Chapter

dpunkt.verlag | it-governance.dpunkt.de | isaca.de

Sonderdruck
aus:

IT-Governance
Fachzeitschrift des ISACA Germany Chapter e.V.

http://it-governance.dpunkt.de/

17. Jahrgang – Heft 38 – Dezember 2023
Seiten 22–27

© dpunkt.verlag GmbH
ISSN 1864-6557

Zum Thema

Nach Microsoft-Schlüssel-Diebstahl aus Crash-Dump – neues IT-Audit-Prüfthema: IT-Diagnose-Daten

Stephen Fedtke

Im Mai 2023 wurde bei Microsoft von chinesischen Hackern der Gruppe Storm-0558 ein Microsoft Master Key aus einem Crash-Dump gestohlen. Spätestens seit diesem Vorfall sollte dieses Sicherheitsrisiko Prüfthema von IT-Revisoren und Auditoren werden. Eigentlich dienen IT-Diagnose-Daten der Fehler- und Problemanalyse. Jede operative IT erzeugt sie fortlaufend. Dennoch enthalten Dumps neben den technischen Daten auch Security-Informationen, Firmengeheimnisse und sogar personenbezogene Daten. Gelangen sie ungeschützt in die Hände Dritter, können sie Angriffe auf die IT-Sicherheit oder Verletzungen des Datenschutzes auslösen.

1 Problemstellung

Wenn Computer oder Applikationen fehlerbedingt abstürzen, halten sie alle Speicherinhalte dieses Vorgangs in sogenannten **Dumps** fest. Der Speicherauszug kann auf dem Monitor, Drucker oder als Datei ausgegeben werden. Die Inhalte können zum Beispiel im Zeichenformat, hexadezimal, binär oder oktal dargestellt werden. Neben der Fehlerdiagnose bei einem Programmabsturz (engl. *crash dump*) wird ein Dump manchmal auch zur Analyse der Leistung genutzt (vgl. [Wikipedia 2022]). Dies bedeutet, dass nicht nur technische Daten in diesem Speicherauszug enthalten sind, sondern auch alle anderen Daten, die sich im Moment des Computerabsturzes im Speicher befinden. Dies können zum Beispiel auch sicherheitsrelevante oder personenbezogene Daten sein. Im Rahmen der Fehlersuche und des Problemmanagements werden diese binär codierten, oftmals Gigabyte großen Dateien mit sehr speziellen Werkzeugen ausgewertet (vgl. [Microsoft 2023]). Meistens sind dies die jeweiligen Debugger des Betriebssystems oder der Entwicklungsumgebung. Sie kennen die internen Strukturen der binär codierten Dump-Files. Aufgrund der technischen Präsentation der Daten ist das Auswertungsverfahren etwas aufwendiger, im Fall textcodierter Daten aber trotzdem einfach. Um einen ersten schnellen Eindruck über die textbezogenen Inhalte von binären Dateien zu erhalten, reichen die Linux- bzw. Unix-Kommandos »string« und »grep« aus (vgl. [Harsh 2023]).

Logs, also Protokolldaten, hingegen können auch manuell eingesehen und analysiert werden. Das thematische Spektrum von Logdateien reicht von applikationsspezifischen Logs bis hin zu standardisierten Systemlog- und Serverlog-Formaten. Sie werden fortlaufend erzeugt und üblicherweise archiviert. In Logfiles befinden sich unter anderem IP-Adressen und User-IDs. Personenbezogene Daten, wie User-IDs, Mitarbeiter- und Kundennamen, sind oft Teil einer technischen Meldung. Ein Beispiel wäre die Meldung zur Festhaltung eines fehlerhaften Logins. Die systematischen Auswertungen von Logs finden üblicherweise toolbasiert statt. Diese Analysen sind wesentlich einfacher durchzuführen als diejenigen von Dump- oder Trace-Dateien. Im Vergleich zu Dumps liegen die Daten in Logs strukturierter vor und können mit einfachen Standardeditoren eingesehen werden, wie z. B. Notepad.

Der Netzwerkverkehr, d. h. IP-Adressen von Sender und Empfänger sowie die eigentlich übertragenen Daten, die sogenannte »Payload«, wird bei Bedarf durch **Traces** mitgeschnitten. Erfolgt die Kommunikation unverschlüsselt, ermöglicht dies eine direkte Einsicht in die Daten. Bestandteil der Payload können neben Nutzdaten auch Kommunikationsdaten wie u. a. Keys und Zertifikate sein. Analog zu Dumps sind auch Trace-Dateien binär codiert und damit nicht so leicht einsehbar wie Logfiles. Ein bekanntes Open-Source-Werkzeug zur Auswertung von Netzwerk-Traces ist z. B. Wireshark.

Insgesamt betrachtet entstehen auf lokalen Servern des unternehmensinternen IT-Betriebs oder auf Systemen der IT-Dienstleister permanent große Mengen an IT-Diagnose-Daten unterschiedlicher Formate und Zwecke. Darüber hinaus entstehen **IT-Diagnose-Daten auch auf IT-basierten elektronischen Geräten**. Im Gesundheitswesen kommt es beispielsweise auf computergestützten Medizingeräten, wie einem CT, MRT oder Röntgengerät, auch zu Funktionsstörungen. Deren IT-Diagnose-Daten müssen ebenfalls für die Problemanalyse vom Hersteller analysiert werden. Prinzipiell stellt auch das »**Internet of Things**« (**IoT**) eine wichtige Quelle für Log- und Diagnosedaten dar.

2 Risikopotenzial

Dumps erwecken auf den ersten Blick den Eindruck, rein technische Daten zu enthalten. Nicht offensichtlich und oft komplett unbekannt ist, dass sie neben den technischen Details auch sehr hohe Volumen **sicherheitskritischer und personenbezogener Daten enthalten**. Diese wahllos eingestreuten sensiblen Daten geraten »per Zufall« in die Dumps, da sie sich im Moment des Computerabsturzes ebenfalls im Speicher befanden. Das Festhalten der Speicherinhalte im Moment eines Abbruchs stellt keine strukturierte Speicherung der verarbeiteten Daten dar, sondern es werden in möglichst kurzer Zeit einfach Memory-Segmente bzw. Pages kopiert.

Darin sind die sensiblen Datenblöcke »wahllos eingestreut und verteilt«. Die breite Streuung und wahllose Verteilung bezieht sich weniger auf die einzelnen Datenelemente bzw. -felder einer größeren Datenstruktur, z.B. eines verarbeitenden Datensatzes, als auf die Datensätze selbst. Durch die lokale Nähe der Einzelteile lässt sich leicht ein inhaltlicher Zusammenhang herstellen.

Insbesondere Dumps abgestürzter Datenbanken und Serverapplikationen erfassen Gigabytes solcher Daten. Sie werden zur Laufzeit, zwecks Verarbeitung und performanter Vorhaltung (»Cache«), im virtuellen und realen Speicher gehalten. Die technischen und sensiblen Daten befinden sich in den IT-Diagnose-Dateien gemäß meiner Erfahrung meist in einem Verhältnis von 50:50. Dies führt zu einem beträchtlichen Umfang an sicherheitskritischen und DSGVO-relevanten, personenbezogenen Daten in einem einzelnen Dump. Beispiele für sicherheitsrelevante Daten wären Passwörter, Keys, Zertifikate oder IP-Adressen. Je nach Anwendung zählen zu den sensiblen Personendaten Informationen über Kunden und Mitarbeiter. Im Gesundheitsumfeld sind es Patienten- und Sozialdaten, bei Behörden entsprechend Bürgerdaten.

Der IT-Betrieb transferiert die IT-Diagnose-Daten regelmäßig per Upload zu den Supportzentren der Softwarehersteller, meistens im Kontext einer Supportanfrage, d.h. eines Problemtickets. Die Destinationen sind u.a. Europa, USA, China oder Indien. Mit im Gepäck reisen die sensiblen Daten. Diese stehen allerdings unter strengem Daten- und Sicherheitsschutz. Umso schlimmer, dass sie einsehbar sind und man nicht weiß, ob, von wem und für welche Zwecke sie ausgewertet werden. Auch macht die Verschlüsselung die IT-Diagnose-Daten nicht sicherer. Denn für das Problem-Handling müssen sie entschlüsselt, verarbeitet und gespeichert werden. Alle damit befassten IT-Experten haben Einblick in die Daten, auch in die sensiblen. Oftmals ist es ein weltweit verteilter Personenkreis aus vielen internen und externen Mitarbeitern.

IT-Diagnose-Daten sind außerdem ein gefährliches IT-Sicherheitsrisiko. Denn sicherheitskritische Informationen über die eigenen IT-Systeme verlassen mit dem Upload nicht anonymisierter IT-Diagnose-Dateien das Haus (Keys, Passwörter etc.). Diese können aus den Dateien bewusst extrahiert und z.B. für einen gezielten Angriff genutzt werden. Dumps und Logs gehören zu den anerkannten **Sicherheitsrisiken gemäß CWE von MITRE**, u.a. dokumentiert in CWE-200 (vgl. [CWE 2023a]) und CWE-528. Zum Beispiel klassifiziert CWE-528 bereits »herumliegende Dump-Dateien« als Risiko: »The product generates a core dump file in a directory, archive, or other resource that is stored, transferred, or otherwise made accessible to unauthorized actors« [CWE 2023b]. CWE-200 war 2021 sogar in den Top 25 Most Dangerous Software Weaknesses als schwerwiegende Sicherheitslücke gelistet (vgl. [CWE 2021]). Die Definition der Top 25 lautet: »This list demonstrates the currently most common and impactful software weaknesses. Often easy to find and exploit, these can lead to exploitable vulnerabilities that allow adversaries to completely take over a system, steal data, or prevent applications from working« [CWE 2023c].

3 Sichtbarkeit kritischer Daten in Dumps

Risikorelevante Daten können mit entsprechenden Werkzeugen sichtbar gemacht werden. Die für die Suche verwendeten Verfahren können grundsätzlich nach der Codierung unterschieden werden:

- textcodierte Daten, bei denen die Vielfalt sogenannter Encodings (utf8, ASCII etc.) eine große Rolle spielt, und
- binär codierte Daten, bei denen die typischen Datenstrukturen (Header, Eyecatcher, IDs etc.) als Ankerpunkte für die Identifikation binärer Details dienen.

Während die meisten DSGVO-relevanten Daten textcodiert sind, liegen die sicherheitsrelevanten Daten, wie z.B. Keys und Zertifikate, meist binär codiert vor. Um diese Datentypen sichtbar zu machen, kommen mehrstufige, die Signifikanz schrittweise erhöhende Verfeinerungs- und Validierungsstufen zur Anwendung. Ziel ist der Ausschluss sogenannter False Positives. Im Fall binärer Details wird die Herausforderung unmittelbar deutlich: »Irgendeine« Folge von 4 Bytes an irgendeiner Stelle im Dump könnte rein theoretisch eine IP-V4-Adresse darstellen.

Prinzipiell sichtbar machen kann man die zufällig verteilten sensiblen Daten innerhalb der binär codierten Dump-Datei mit dem »Klassiker« unter den Forensik-Tools, dem sogenannten Hex-Editor. Abbildung 1 zeigt, wie ein Stück sensibler Daten im Dump arrangiert sein kann. Im Übrigen, Standardtexteditoren, wie z.B. Notepad oder Word, sind ungeeignet für die Inspektion von Dumps und Traces. Die binäre Codierung führt auch bei der Anonymisierung zu einer direkten Herausforderung: die Gefahr, dass durch »zu großzügige Änderungen« eine korrupte Datei entsteht bzw. die technische Aussagekraft der IT-Diagnose-Dokumente infrage gestellt wird. Im Endeffekt reicht ein falsches Bit.

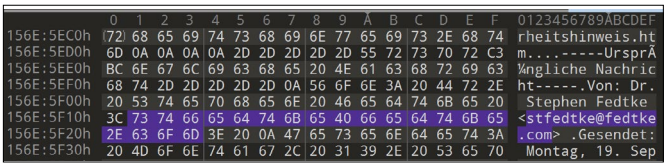

Abb. 1: Anzeige eines Dump-File-Ausschnitts mit eingelagerten personenbezogenen Textdaten im Hex-Editor

Hacker können bei einem Angriff direkt auf abgespeicherte, ungeschützte Dumps zugreifen und die darin enthaltenen sensiblen Informationen extrahieren und missbrauchen. Das zeigt der Vorgang des Schlüsseldiebstahls aus dem Crash-Dump im Microsoft Debug-Bereich laut heise: »Da der Crash-Dump vermeintlich keine sensiblen Daten enthielt, wurde er aus der isolierten Produktionsumgebung in die Debug-Umgebung im internetverbundenen Unternehmensnetz verschoben – ein Standardvorgang, betont Microsoft.

Ein bösartiger Akteur der Storm-0558-Gruppe habe danach einen Unternehmensnetzzugang eines Microsoft-Ingenieurs kompromittiert. Das Konto hatte Zugriff auf die Debug-Umgebung mit dem Crash-Dump, der fälschlicherweise den Schlüssel enthielt« [Knop 2023]. Wie später bekannt wurde, sind mit diesem Azure Cloud Signing Key 60.000 Mails aus zehn Accounts sowie Listen mit allen Mailadressen des State Departments gestohlen worden. Auch europäische Behörden waren betroffen (vgl. [Steevens 2023]). Durch diesen Master-Key-Diebstahl bei Microsoft wird sich aus Autorensicht das Ranking von CWE-528 und CWE-200 im Jahr 2024 sicher nach oben verändern. Für alle Unternehmen, Behörden und insbesondere für sogenannte kritische Infrastrukturen (KRITIS) ist ein solcher Vorfall definitiv ein inakzeptables Risiko. Auch können Verfechter der ZeroTrust-Idee diese Sicherheitslücke nicht riskieren. Dies gilt auch vor dem Hintergrund, dass nicht nur Hacker, sondern auch die Mitarbeiter der Softwarehersteller gezielte Datenextraktionen vornehmen können.

4 Sicherer Umgang mit IT-Diagnose-Daten

Die Anonymisierung sämtlicher in IT-Diagnose-Daten enthaltenen sensiblen Daten bietet Schutz vor Cybercrime und Verletzungen der Datenschutz-Grundverordnung (DSGVO). Sie befreit Dumps, Logs und Traces von datenschutz- und sicherheitskritischen Inhalten vor dem Upload und lässt die technischen Daten unberührt. Erst nach dieser Bereinigung werden die Dateien zum Support des Softwareherstellers hochgeladen – risikominimiert und revisionsgerecht. Durch Begleitdokumente kann man den Gesamtprozess transparent und nachweisbar für Revision und Datenschutz machen. Im Gegensatz zu vergleichbaren Anwendungsfällen, wie z. B. der Testdatengenerierung, benötigt man für die Anonymisierung von IT-Diagnose-Dokumenten gezielt ausgerichtete Algorithmen und Verfahren.

Alternativen zur Anonymisierung wären: a) die grundsätzliche Vermeidung und Minimierung von IT-Diagnose-Daten, b) die Nachstellung eines Problems auf unkritischen, isolierten Testsystemen und c) der Remote-Zugriff vonseiten des externen Supports. Sämtliche Alternativen zeichnen sich durch einen erheblichen Mehraufwand, hohe Komplexität und teilweise Zweckverfehlung aus. So würde zum Beispiel die Nachstellung von Abstürzen bewährter professioneller Software in höheren Versionen oft hochspezifische Sondersituationen bedingen. Deren Nachstellen, z. B. mit Testdaten, wäre sehr aufwendig und mit dem üblichen Arbeitsdruck und Mangel an Spezialisten schwer zu rechtfertigen.

Eine weitere Möglichkeit wäre ein Remote-Zugriff der Hersteller verbunden mit einer lokalen Analyse. Üblicherweise steht diese »VIP-Behandlung« aufseiten der Hersteller aus Kapazitäts- und Kostengründen nur großen Rechenzentren mit Premium-Supportvertrag zur Verfügung. Aber auch eine solche umgekehrte Zugriffsrichtung auf die IT-Diagnose-Daten unterliegt denselben strengen Datenschutzregelungen. Entsprechend muss ein uneingeschränkter Remote-Zugriff für Hersteller auf die IT-Systeme und Daten des Kunden formell durch Policies und technisch durch eine geeignete Konfiguration der Systeme begrenzt und kontrolliert werden. Ganz zu schweigen von einem automatischen und damit unkontrollierten Transfer von IT-Diagnose-Daten an die Hersteller, einem absoluten »No-Go«.

5 Lebenszyklus von IT-Diagnose-Daten

5.1 Überblick

Das Risiko um IT-Diagnose-Daten entsteht beim Upload-Vorgang, d. h., sämtliche mit IT-Diagnose-Daten einhergehenden Gefahren entstehen in dem Moment, da Dumps, Logs und Traces das Haus verlassen. Eine präzise Risikoanalyse muss entlang des gesamten Lebenszyklus der IT-Diagnose-Daten verlaufen. Die typische praktische Zusammenarbeit von Softwareanwender und -hersteller ist in Abbildung 2 dargestellt.

Wie aufgezeigt, teilt sich der zu analysierende Prozess in zwei grundsätzliche Bereiche bzw. Phasen:

- Der IT-Betrieb ist Ursache und Quelle von IT-Diagnose-Daten und vollzieht den Upload.
- Der Softwarehersteller oder Dienstleister empfängt, speichert und verarbeitet diese Daten.

Das heißt, die genutzte Upload-Funktionalität verkörpert die Schnittstelle zwischen beiden Institutionen. Der Upload erfolgt üblicherweise im Kontext einer Webanwendung für das Problemmanagement, z. B. in einem sogenannten Ticketsystem. Bei kleineren Herstellern oder sehr großen Datenmengen wird der Transfer gerne auch direkt und »einfach so« via (S)FTP vollzogen. Zweifelsohne, eine verschlüsselte Übertragung ist Standard, diese verhindert aber keinesfalls die hier diskutierten Risiken. Denn der Softwaresupport wird die IT-Diagnose-Datei nach dem Empfang entschlüsseln, speichern und verarbeiten.

Nach dem Upload hat der Kunde keinen wirklichen Einblick mehr in die genauen Abläufe aufseiten des Herstellers, in dessen interne Verarbeitungsstrukturen und die involvierten Subunternehmen. Hersteller sprechen hierüber auch nicht offen. Trotzdem bestätigt die langjährige IT-Praxis die im folgenden Abschnitt beschriebenen Tatsachen und Erfahrungen als repräsentativ.

5.2 Übertragung von Diagnosedaten zum Softwarehersteller

Die vom IT-Betrieb versendeten IT-Diagnose-Daten gelangen in mehreren Schritten, und damit über mehr als eine Zwischenspeicherungslokation, zur finalen »zentralen« Speicherlokation. Sämtliche involvierten Speicherlokationen sind weltweit verteilt und der gesamte Weg kann über Subunternehmen und Tochtergesellschaften verlaufen. Jede dieser Lokationen impliziert ein prinzipielles Risiko. Manchmal erfolgt der initiale Upload zuerst zu einer ver-

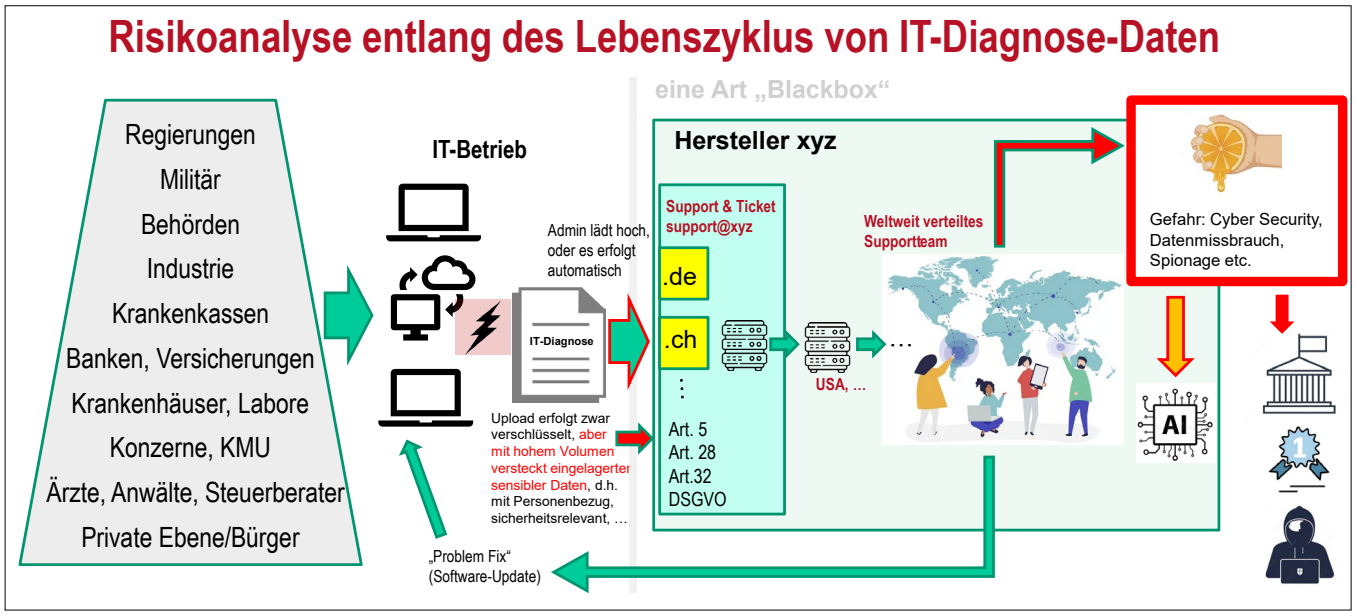

Abb. 2: Lebenszyklus von IT-Diagnose-Daten: Entstehung, Upload, Analyse und mögliche »Sonderverwertungen«

trauensvollen lokalen oder lokal wirkenden Supportadresse. Das heißt, Dumps, Logs und Traces werden z. B. nicht direkt in die USA übertragen, sondern erst einmal von der europäischen oder landesbezogenen Niederlassung empfangen und anschließend intern weitergereicht. Die Support- und Pflegeverträge im Rahmen der Auftragsvereinbarungen großer Hersteller lassen die oft sehr komplexe, weltweite interne Verflechtung und Verkettung im Umgang mit IT-Diagnose-Daten erahnen. Hersteller halten sich meist sämtliche Optionen offen. Aber eines gilt es festzuhalten: Trotz des umfangreichen »Paperwork« bleibt die Übertragung, Speicherung und Verarbeitung der in IT-Diagnose-Daten enthaltenen sensiblen Daten stets zweckfremd und damit juristisch angreifbar. Eine besondere Situation liegt vor, wenn IT-Diagnose-Daten automatisch zum Hersteller übertragen werden, eventuell sogar ohne Kenntnisnahme oder Genehmigung des Kunden. Für professionelle IT-Umgebungen ist dies ein »No-Go« und sollte durch eine Policy oder Systemkonfiguration unterbunden werden.

5.3 Analyse der empfangenen IT-Diagnose-Daten

In einem ersten Schritt werden die empfangenen IT-Diagnose-Dateien entschlüsselt und durch automatisierte, toolbasierte Prozesse voruntersucht. Auch anhand statistischer Auffälligkeiten können grundsätzliche Probleme bereits sehr schnell entdeckt werden. Diese Vorstufe entlastet den First-Level-Support und befreit ihn von Routinearbeiten. Oftmals werden Probleme auf diese Weise sogar »automatisch gelöst«, und zwar dann, wenn das Problem bekannt und eine Lösung bereits vorhanden ist. Ist eine solche schnelle Lösung nicht möglich, greifen Mitarbeiter des First- und Second-Level-Supports im Kontext der Problemticket-Bearbeitung auf die jeweiligen IT-Diagnose-Daten direkt zu. Hierzu müssen diese ggf. auf ihre lokalen Rechner heruntergeladen werden oder die Analysten verarbeiten direkt die zentral gespeicherten Dateien, um eine grundsätzliche Analyse und Abklärung vorzunehmen. Die Praxis zeigt, dass meistens alle Supportmitarbeiter auf alle IT-Diagnose-Daten Zugriff haben. Hierzu zählen Hunderte oder Tausende Personen, intern und extern, weltweit verteilt. Anders ließen sich die organisatorischen Aspekte wie Stellvertretung, Eskalation, 24-Stunden-Bereitschaft, weltweite Abdeckung, Kostenminimierung etc. nicht sinnvoll zusammenführen. Die Entwickler in den Laboren kommen erst dann ins Spiel, wenn das Problem vom First- und Second-Level-Support nicht wirklich analysiert oder behoben werden kann. Dies passiert z. B., wenn detailliertes Codewissen und Zugriff auf die Compiler-Listings bzw. den Sourcecode notwendig sind. Zu den Mitarbeitern des Supports kommen also die Entwickler als weiterer, nicht minder großer Personenkreis hinzu, mit Lesezugriff auf Dumps, Logs und Traces der weltweiten Kunden.

5.4 Gefahr der missbräuchlichen Verwendung von Diagnosedaten

Betrachtet man die oben beschriebene Situation des Supportbetriebs vor dem Hintergrund der hohen Anzahl weltweit verteilter Mitarbeiter, ausgestattet mit Lesezugriff auf die IT-Diagnose-Dateien, ist eine gewisse Sorge berechtigt. Sie lässt sich u. a. in folgender Frage zusammenfassen: Wozu werden die Daten noch verwendet, außer zu der konkreten softwaretechnischen Problemanalyse? Zum einen geht es um mögliche »interne Sonderverwertungen« in Form systematischer Auswertungen von IT-Diagnose-Daten durch den Hersteller. Abgesehen von rein statistischen Auswertungen sind weitere, viel höherwertigere Formen der gezielten Datenextraktion vorstellbar. Hersteller haben den klaren Transparenzvorteil, d. h., sie kennen sämtliche Datenformate und wissen exakt, wo welche Informationen zu finden sind.

Aber auch jeder einzelne der vielen Mitarbeiter kann individuelle Eigeninteressen an den »versteckt« eingelagerten sensiblen Daten entwickeln und diese gezielt extrahieren. Die erlangten Informationen können für verschiedene Zwecke missbraucht werden. Sie können z. B. veröffentlicht, im Darknet verkauft oder für Erpressung oder Angriffe genutzt werden. Ja, Dumps von Produktionssystemen interessanter Unternehmen oder Behörden haben definitiv einen Verkaufswert im Darknet. Der von der Aufgabenstellung im Unternehmen bereits gegebene Lesezugriff auf die Dateien, das vorhandene technische Wissen und eine persönliche Zielsetzung reichen vollkommen aus. Im Übrigen braucht man für ein solches kriminelle Vorgehen keine Unterstützung von Kollegen. Man kann es allein und diskret durchführen. Das ist ein großer Vorteil für potenzielle Akteure. Damit steigt die Risiko-Eintrittswahrscheinlichkeit nochmals erheblich. Eine heiklere Situation tritt ein, wenn offizielle Stellen, z. B. Regierungsbehörden oder Geheimdienste, sich an den Softwarehersteller wenden, um Zugriff auf IT-Diagnose-Daten ausgewählter Unternehmen oder Institutionen zu erhalten. Deren erfahrene Forensiker können gezielt nach Informationen suchen. Dumps aus produktiven Umgebungen sind sehr interessant wegen ihrer großen Mengen an Daten über Kunden, Mandanten oder etwa Patienten. Entsprechend enthalten sie zahllose technische Details über die IT-Infrastruktur, wie z. B. IP-Adressen, Keys, Passwörter oder eingesetzte Sicherheitslösungen. Dumps sind eine wahre »Goldgrube«. Bei Angriffen auf Hersteller können Hacker auch IT-Diagnose-Daten von Kunden finden und entwenden. Diese Klasse von Profis ist sehr geübt in der gezielten Auswertung von Dumps, Logs und Traces.

Im Endeffekt können viele Risikoaspekte der ZeroTrust-Idee auch auf den Bereich der IT-Diagnose-Daten übertragen werden. Die sensiblen Infrastrukturdaten, Systemdetails und Sicherheitsdaten in IT-Diagnose-Dateien bieten ein hohes Verwertungs- bzw. Missbrauchspotenzial. Natürlich stellt sich die folgende Frage: Wie relevant ist das oben dargestellte Risiko für einen IT-Betrieb? Wie oft verlassen Dumps, Logs und Traces eigentlich das Haus und werden zu Herstellern transferiert? Wie hoch ist das geschätzte Risikovolumen? Für einen IT-Betrieb durchschnittlicher Größe sind bis zu 150 Supporttickets pro Jahr mit IT-Diagnose-Daten gängige Praxis. Von diesen sind ca. 75 % aus der IT-Produktion. Laut einer aktuellen Bitkom-Studie (vgl. [Bitkom 2022]) gibt es rund 50.000 Rechenzentren in Deutschland. Eine defensive Überschlagsrechnung ergibt daher: Deutschlandweit werden circa 100 x 50.000 = 5 Millionen IT-Diagnose-Dateien mit sensiblen Daten an Hersteller in die ganze Welt verschickt, pro Jahr.

Entsprechend fragen sich Unternehmen und Behörden: Wie wahrscheinlich ist es, entsprechenden Vorwürfen eines DSGVO-Verstoßes ausgesetzt zu werden? Oder wer würde die Sicherheitslücken durch offen zugängliche Schlüssel, Passwörter oder Zertifikate in IT-Diagnose-Daten monieren? Wer weiß eigentlich von meinen Uploads? Im Endeffekt gibt der vorangehende Punkt bereits eine Antwort: Jede professionell betriebene IT erzeugt IT-Diagnose-Daten und führt zur Kooperation mit Herstellern und Dienstleistern. Das heißt, der Vorwurf und Nachweis eines Verstoßes ist relativ einfach und die juristische Kerngefahr bestünde in einer Beweislastumkehr. Der Betreiber der IT müsste dann nachweisen, keinen Upload vorgenommen zu haben. Oder er müsste die sensiblen Daten schützen, zum Beispiel durch Anonymisierung, nachweislich etwa durch Begleitdokumente.

6 Rechtliche Folgen

Diese Risikobetrachtungen dürfen nicht zu einer Diskreditierung der Hersteller und Supportzentren bzw. deren Mitarbeiter veranlassen. Aber selbst in einem fiktiven Vertrauenskontext vollständiger Sicherheit und Compliance aufseiten des Herstellers stellt eine zweckfreie Weitergabe personenbezogener Daten einen DSGVO-Verstoß dar. Denn die in IT-Diagnose-Dateien versteckt eingelagerten personenbezogenen und sensiblen Daten sind für den eigentlichen Zweck, und zwar für die Lösung des Softwareproblems, in keiner Weise notwendig. So äußert sich der Datenschutzexperte RA Horst Speichert: »Die genannten Datenexporte enthalten standardmäßig ›Überhänge‹, also ›zu viele‹ Daten, die gemäß den Datenschutzanforderungen wegen fehlendem ›Zweck‹ und ›Notwendigkeit‹ nicht an Dritte weitergeleitet werden dürfen. Diese in ihrer Struktur ›mehrdimensionalen Überhänge‹, in denen sich auch hoch-sensible Personendaten befinden, sind oftmals unverhältnismäßig und demnach rechtswidrig« [Speichert 2023]. Aus DSGVO-Perspektive erfüllen sie nicht das Erfordernis von Zweck und Nutzen für einen Versand an Dritte und Drittländer. Ein solcher Datentransfer verstößt gegen das europäische und das deutsche Datenschutzgesetz. Das heißt, jegliches »Philosophieren« darüber, ob es kriminelle Energie aufseiten des Herstellersupports geben könnte, ist nicht wirklich relevant. Denn selbst in einer idealen Welt führt der zweckfreie Transfer direkt zum Thema DSGVO-Verstoß.

Die datenschutzrechtliche Position ist eindeutig und reklamiert den Schutz der personenbezogenen Daten in Dumps, Logs und Traces per se. Compliance und IT-Governance richten ihr Augenmerk auf alle Risiken, die den IT-Betrieb und somit die Unternehmensprozesse negativ beeinträchtigen könnten. Dies wäre bei einem Diebstahl von sicherheitskritischen Informationen, wie etwa von Softwareschlüsseln oder Passwörtern aus IT-Diagnose-Daten oder Traces, der Fall. Der bloße Versand von IT-Diagnose-Daten mit personenbezogenen Daten an Dritte ist per se eine Datenschutzverletzung. Sicherheitsbedrohungen hingegen entstehen erst durch aktives, kriminelles Handeln Dritter, wie beispielsweise die gezielte Extraktion von sicherheitsrelevanten Daten und deren Verkauf oder Missbrauch, z. B. für einen Angriff.

Outsourcing und Cloud-Nutzung verlagern das Problem und führen zur Einbindung weiterer Parteien und insgesamt noch komplexeren Verantwortlichkeiten. Denn die IT-Diagnose-Daten entstehen jetzt auf Systemen, deren Betrieb und Problemmanagement nicht mehr in den eigenen Händen liegt.

Die operativen Teams der Dienstleister entscheiden beim Upload der IT-Diagnose-Daten u.a. über die datenschutzrelevanten Daten ihrer Auftraggeber, d.h. über fremde Daten.

Cyberversicherungen haben das Risiko bereits erkannt. Der Upload der IT-Diagnose-Dateien ist ein von den IT-Betriebs-Spezialisten bewusst vollzogener Arbeitsschritt. Man könnte sozusagen einen gewissen Vorsatz unterstellen. Daraus ergeben sich zweierlei Risiken: mögliche DSGVO-Bußen und Schadensersatzansprüche. Letztere wurden in Bezug auf die DSGVO-bezogene Haftung mit dem EuGH-Urteil vom 4.5.2023 nochmals präzisiert. Juristische Akteure könnten es mit der gefürchteten Beweislastumkehr strategisch verbinden. Underwriter bei Cyberversicherungen ergänzen die Anonymisierung in den Obliegenheiten.

7 Zusammenfassung

Datenschützer, IT-Sicherheitsverantwortliche, IT-Revisoren und IT-Auditoren sind gleichermaßen gefragt. IT-Diagnose-Daten haben mit der DSGVO den Bereich eines Schattendaseins verlassen. Durch den Microsoft-Hack und den massiven Angriff auf die Datensicherheit erfahren sie aktuell erhöhte Aufmerksamkeit in der IT-Sicherheit. Durch die Risiken für Unternehmen und Behörden und deren Verantwortliche ergeben sich für Auditing, Revision und Datenschutz neue Prüfbereiche der IT-Security und Compliance.

Insgesamt wird deutlich, dass ein nicht sicherheits- und rechtskonformer Umgang mit IT-Diagnose-Daten zum Unternehmensrisiko wird und folglich revisionsrelevant ist. Die Risiken des freien Versands von IT-Diagnose- und -Protokolldaten sind vielfältig. Sie reichen von der Verletzung rechtlicher Auflagen, wie insbesondere des Datenschutzes, über den unkontrollierten Verlust von Firmengeheimnissen bis hin zur Offenlegung sicherheitskritischer Details. Diese können Angreifern zum schnellen Erfolg verhelfen, wie jüngst bei Microsoft durch die Extraktion eines Master Key passiert. Dagegen sind Schutzmaßnahmen, wie etwa die automatisierte Anonymisierung der datenschutz- und sicherheitsrelevanten Daten in IT-Diagnose-Dateien, wenig aufwendig. IT-Auditoren und IT-Revisoren sind, zusammen mit IT-Sicherheit und Datenschutz, gefordert, den IT-Betrieb zu instruieren und zu überprüfen (vgl. [Fedtke 2023]). Die Einbeziehung von IT-Diagnose-Daten in ihre Prüfschemata erfüllen die Anforderungen an Unternehmenssicherheit und Compliance.

Literatur

[Bitkom 2022] *Bitkom e.V. (Hrsg.):* Rechenzentren in Deutschland – Aktuelle Marktentwicklungen, Stand 2022. Borderstep Institut, Berlin, 2022, S. 8; *https://www.bitkom.org/sites/main/files/2022-02/10.02.22-studie-rechenzentren.pdf.*

[CWE 2021] *Common Weakness Enumeration:* CWE VIEW: Weaknesses in the 2021 CWE Top 25 Most Dangerous Software Weaknesses, 2021; *https://cwe.mitre.org/data/slices/1337.html;* Zugriff am 17.10.2023.

[CWE 2023a] *Common Weakness Enumeration:* CWE-200: Exposure of sensitive Information to an Unauthorized Actor, 2023; *https://cwe.mitre.org/data/definitions/200.html*; Zugriff am 17.10.2023.

[CWE 2023b] *Common Weakness Enumeration:* CWE-528: Exposure of Core Dump File to an Unauthorized Control Sphere, 2023; *https://cwe.mitre.org/data/definitions/528.html*; Zugriff am 17.10.2023.

[CWE 2023c] *Common Weakness Enumeration:* CWE Top 25 Most Dangerous Software Weaknesses, 2023*; https://cwe.mitre.org/index.html*; Zugriff am 17.10.2023.

[Fedtke 2023] *Fedtke, S.:* Revision, Revisionsprüfung. In: Sicherheitsrisiko durch datenschutzwidrigen Abfluss sensibler Daten und Geschäftsgeheimnisse im IT-Betrieb. Whitepaper, 2023, S. 9*; https://www.enterprise-it-security.com/IT-Diagnose-Daten-Whitepaper-DEV4.pdf#page=9.*

[Harsh 2023] *Harsh, A.:* How to Search Strings in Binary Files, 2023. In: LinuxOPsys; *https://linuxopsys.com/topics/search-strings-in-binary-files*; Zugriff am 17.10.2023.

[Knop 2023] *Knop, D.:* Gestohlener Microsoft-Schlüssel stammte aus einem Crash-Dump, 2023, heise online; *https://www.heise.de/news/Gestohlener-Microsoft-Schluessel-stammte-aus-einem-Crash-Dump-9297240.html*; Zugriff am 17.10.2023.

[Microsoft 2023] *Microsoft:* Analysieren von Absturzabbilddateien mithilfe von WinDbg, 2023. Microsoft Learn; *https://learn.microsoft.com/de-de/windows-hardware/drivers/debugger/crash-dump-files*; Zugriff am 17.10.2023.

[Speichert 2023] *Speichert, H.:* DSGVO-Haftungs- und Sicherheitsrisiken durch Protokoll- und Diagnosedaten im IT-Betrieb. Datenschutz und Datensicherheit (DuD) 47, 04/2023, S. 229; *https://www.enterprise-it-security.com/DuD-Artikel-Speichert-042023/GOV.*

[Steevens 2023] *Steevens, P.:* 60.000 geklaute Regierungsmails: Erste Zahlen nach Microsofts Cloud-Key-Debakel, 2023, heise online; *https://www.heise.de/news/60-000-geklaute-Regierungsmails-Erste-Zahlen-nach-Microsofts-Cloud-Key-Debakel-9321044.html?wt_mc=nl.red.security.security-nl.2023-10-02.link.link*; Zugriff am 17.10.2023.

[Wikipedia 2022] *Wikipedia:* Dump, *2022; https://de.wikipedia.org/wiki/Dump*; Zugriff am 17.10.2023.

Stephen Fedtke

ist Wirtschaftsingenieur und CTO von ENTERPRISE-IT-SECURITY.COM, ein auf IT-Sicherheit und Compliance-Lösungen spezialisierter Dienstleistungsbereich des Unternehmens Dr. Stephen Fedtke System Software. Außerdem ist er Autor und Herausgeber im Verlag Springer Vieweg.

Dr. Stephen Fedtke
Enterprise-IT-Security.com
Seestr. 3a
CH-6300 Zug
Schweiz
stfedtke@enterprise-it-security.com
www.enterprise-it-security.com