Holger Reibold

KI Sicherheit

Einstieg in die Praxis

Angriffe

Verteidigungsstrategien

Sicherheitsarchitekturen

BRAIN-MEDIA.DE

Holger Reibold

KI-Sicherheit

Einstieg in die Praxis

Angriffe, Risiken,

Verteidigungsstrategien

BRAIN-MEDIA.DE

ISBN: 978-3-95444-302-4

Cover: Freepik

Brain-Media.de
Dr. Holger Reibold – Hubert-Müller-Str. 52c – 66113 Saarbrücken

info@brain-media.de – www.brain-media.de

Inhaltsverzeichnis

Vorwort

Die Geschichte der Informationssicherheit ist immer auch die Geschichte ihrer Angreifer. In den frühen Jahrzehnten der Computerära waren Hacker Menschen, die Systeme durch geschickte Manipulation von Code, Protokollen oder Hardware kompromittierten. Sie saßen an Terminals, analysierten Binärdateien, schrieben Exploits, überlisteten Authentifizierungsmechanismen und nutzten Schwachstellen in Betriebssystemen und Netzwerken aus.

Dieses Bild prägt bis heute unser Sicherheitsdenken. Firewalls, Intrusion-Detection-Systeme, Patch-Management und Penetrationstests beruhen auf der Annahme, dass Angriffe primär auf Code, Speicher und Schnittstellen abzielen.

Doch dieses Modell greift nicht mehr.

Mit dem Aufstieg moderner KI-Systeme ist ein neues Bedrohungsprofil entstanden – eines, das nicht zwingend menschlich ist, keine Tastatur benötigt und keinen klassischen Exploit schreibt. Angriffe richten sich nicht mehr primär gegen Softwarekomponenten, sondern gegen Modelle, Daten und Entscheidungsprozesse.

Der Angreifer greift nicht den Computer an.

Er greift die Wahrnehmung der Maschine an.

Seit dem Durchbruch des Deep Learning Mitte der 2000er-Jahre hat sich künstliche Intelligenz von einem Forschungsthema zu einer produktiven Schlüsseltechnologie entwickelt. KI-Systeme übertreffen Menschen in der Bilderkennung, analysieren medizinische Befunde, steuern Fahrzeuge und unterstützen komplexe Entscheidungsprozesse in Wirtschaft und Verwaltung.

Gleichzeitig hat sich gezeigt: Diese Systeme sind nicht robust im klassischen sicherheitstechnischen Sinne. Ihre Leistungsfähigkeit beruht auf statistischer Generalisierung, nicht auf formaler Korrektheit oder deterministischem Verhalten.

Forschungsarbeiten wie die von Hu et al. kommen zu einem nüchternen, aber folgenreichen Befund: KI-Systeme sind durchaus intelligent, aber fragil.

Diese Fragilität ist kein Implementierungsfehler einzelner Modelle, sondern eine systemische Eigenschaft. KI-Systeme sind über ihren gesamten Lebenszyklus hinweg angreifbar:

- Datensammlung: Verzerrte, manipulierte oder absichtlich gefälschte Daten
- Vorverarbeitung: Skalierungs- und Normalisierungsangriffe
- Training: Data Poisoning, Backdoors, Modellmanipulation
- Inferenz: Adversarial Inputs, Kontextmanipulation, Prompt Injection
- Systemintegration: API-Leaks, fehlerhafte Zugriffskontrollen, Code-Schwachstellen

Jede dieser Phasen eröffnet Angriffsflächen, die sich mit klassischen Sicherheitsmechanismen kaum erfassen lassen.

Die Angriffe auf KI-Systeme folgen nicht den vertrauten Paradigmen der IT-Sicherheit. Statt Buffer Overflows oder Code Injection nutzen Angreifer mathematische Eigenschaften von Lern- und Optimierungsverfahren, Gradienten- und Wahrscheinlichkeitsinformationen, minimale, für Menschen kaum wahrnehmbare Eingabeveränderungen sowie visuell oder semantisch unauffällige Trigger in Trainings- und Kontextdaten. Ein einzelnes manipuliertes Bild, ein gezielt formulierter Texteingang oder ein präparierter Datensatz kann ausreichen, um ein System zuverlässig fehlzuleiten – ohne eine einzige Zeile Code zu verändern.

Die Konsequenzen sind real. Studien dokumentieren schwerwiegende Fehlentscheidungen in sicherheitskritischen Systemen. Zwischen 2000 und 2013 wurden allein in den USA über hundert Todesfälle im Zusammenhang mit robotisch assistierten chirurgischen Systemen gemeldet. In vielen Fällen spielten Fehlinterpretationen sensorischer oder bildbasierter Eingaben eine entscheidende Rolle. Diese Vorfälle sind keine Ausnahmen. Sie sind frühe Symptome eines Problems, das mit der zunehmenden Integration von KI-Systemen weiter eskaliert.

In diesem Kontext stellt sich die Frage, warum klassische Sicherheitskonzepte versagen. Firewalls, Penetrationstests und Compliance-Frameworks sind darauf ausgelegt, Systeme zu schützen. KI-Sicherheit hingegen muss Entscheidungen, Kontexte und Wahrnehmungsketten absichern.

Ein korrekt gepatchtes, vollständig gehärtetes System kann dennoch katastrophale Fehlentscheidungen treffen, wenn Trainingsdaten manipuliert wurden, Kontextinformationen verfälscht sind oder Eingaben gezielt adversarial gestaltet wurden. Diese Angriffe hinterlassen keine klassischen Spuren. Sie tauchen in Logs nicht als „Attacke“ auf, sondern als scheinbar legitime Nutzung des Systems.

Genau hier setzt dieses Buch an. Dieses Buch ist kein Überblick über KI-Forschung und kein ethisches Diskussionspapier. Es richtet sich an Praktikerinnen und Praktiker, die KI-Systeme entwerfen, betreiben und absichern müssen. Sie lernen:

- wo reale Angriffsflächen moderner KI-Systeme liegen
- wie diese Angriffe praktisch funktionieren
- warum etablierte Sicherheitsmaßnahmen oft wirkungslos bleiben
- und wie sich KI-Systeme mit offenen Werkzeugen und soliden Architekturprinzipien absichern lassen

KI-Sicherheit ist keine akademische Disziplin mehr. Nein: Sie ist eine operative Notwendigkeit.

Herzlichst

Holger Reibold

(Januar 2026)

1 Warum klassische IT-Sicherheit bei KI-Systemen versagt

Die klassische IT-Sicherheit beruht auf einem vergleichsweise stabilen Weltbild. Systeme bestehen aus klar definierten Komponenten: Betriebssysteme, Anwendungen, Netzwerke und Schnittstellen. Sicherheit bedeutet, diese Komponenten gegen unautorisierten Zugriff, Manipulation und Ausfall zu schützen. Wenn der Code korrekt ist, Konfigurationen stimmen und bekannte Schwachstellen geschlossen sind, gilt ein System als hinreichend sicher für den produktiven Einsatz.

1.1 Sicherheit unter falschen Annahmen

Doch das klassische Sicherheitsmodell ist über Jahrzehnte gewachsen – und war für klassische Software erfolgreich. Es setzt voraus, dass sich das Verhalten eines Systems aus seinem Code ableiten lässt und dass sicherheitsrelevante Fehler sich als technische Abweichungen manifestieren: Speicherfehler, Logikfehler, fehlerhafte Zugriffskontrollen, unsichere Protokolle.

KI-Systeme entziehen sich diesem Modell.

Nicht, weil sie „unsicher programmiert“ wären, sondern weil ihr Verhalten nicht mehr primär durch explizite Regeln bestimmt wird. Moderne KI-Systeme lernen aus Daten. Sie approximieren Funktionen,

erkennen Muster und treffen Entscheidungen probabilistisch. Sicherheit bezieht sich damit nicht mehr nur auf das System, sondern auf das Verhalten, das dieses System unter realen Bedingungen zeigt.

Das ist ein fundamentaler Bruch mit dem bisherigen Sicherheitsdenken.

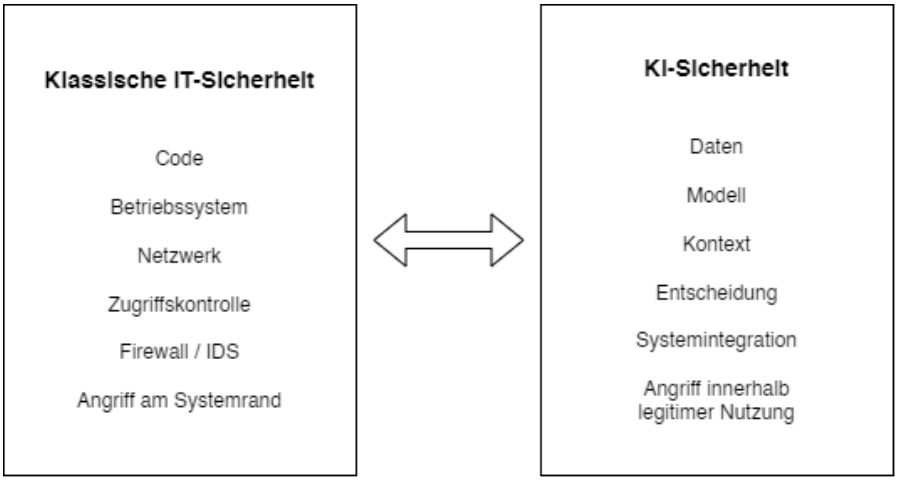
1.2 Der Verlust des klaren Systemrandes

Ein zentrales Konzept der IT-Sicherheit ist der Systemrand. Es gibt ein „Innen" und ein „Außen", legitime Nutzer und Angreifer, erlaubte und unerlaubte Aktionen. Sicherheitsmaßnahmen wie Firewalls, Authentifizierung oder Netzwerksegmentierung sind darauf ausgelegt, diesen Rand zu kontrollieren. KI-Systeme verwischen diesen Rand.

Ein Angreifer muss kein Netzwerksegment überwinden, keine Authentifizierung umgehen und keine privilegierten Zugriffe erlangen. In vielen Fällen genügt es, mit dem System auf genau die Weise zu interagieren, für die es vorgesehen ist. Eingaben erfolgen über offizielle Schnittstellen, Formate und Protokolle. Aus Sicht der Infrastruktur ist alles korrekt. Der Angriff findet nicht am Rand statt, sondern innerhalb der vorgesehenen Nutzung.

Ein Sprachmodell, das durch gezielte Eingaben dazu gebracht wird, interne Informationen preiszugeben, wird nicht kompromittiert. Es erfüllt seine Funktion – nur auf eine Weise, die den Sicherheitsannahmen der Betreiber widerspricht. Ein Bilderkennungssystem, das manipulierte Eingaben falsch klassifiziert, verarbeitet valide Bilddaten. Ein Empfehlungssystem, das systematisch verzerrte Entscheidungen trifft, reagiert

auf legitime Nutzersignale. Damit verliert die klassische Trennung zwischen Nutzung und Angriff ihre Bedeutung.



Klassische IT-Sicherheit basiert auf deterministischen Systemen mit klaren Fehlzuständen und behebbarem Fehlverhalten. KI-Sicherheit muss mit probabilistischem, kontextabhängigem Verhalten umgehen, bei dem Risiken nicht verhindert, sondern nur begrenzt werden können.

1.3 Sicherheit ohne Exploit

In der traditionellen IT-Security ist der Exploit das zentrale Objekt. Ein Exploit nutzt einen Fehler in der Implementierung oder Konfiguration aus, um Kontrolle zu erlangen oder Schutzmechanismen zu umgehen. Die gesamte Toollandschaft – Scanner, IDS, SIEM, WAFs – ist darauf

ausgerichtet, solche Muster zu erkennen. Viele Angriffe auf KI-Systeme kommen ohne Exploit aus.

Sie nutzen keine Programmierfehler, sondern Eigenschaften, die für das Funktionieren der Modelle notwendig sind. Generalisierung bedeutet, dass ein Modell auf unbekannte Eingaben reagieren kann. Toleranz gegenüber Rauschen erlaubt robuste Verarbeitung realer Daten. Kontextabhängigkeit ermöglicht flexible Interpretation. Genau diese Eigenschaften lassen sich gezielt ausnutzen.

Ein adversariales Beispiel verletzt keine Sicherheitsrichtlinie. Es ist eine valide Eingabe, die statistisch so konstruiert ist, dass das Modell sie falsch interpretiert. Prompt-Injection ist kein Code-Injection-Angriff, sondern eine semantische Manipulation des Kontexts. Data Poisoning verändert nicht den Code, sondern die Erfahrungsbasis des Modells. Aus Sicht klassischer Sicherheitsmodelle passiert nichts Illegitimes. Aus Sicht der Anwendung entsteht ein sicherheitsrelevanter Fehler.

1.4 Wenn korrekter Betrieb falsche Ergebnisse liefert

Ein besonders gefährlicher Aspekt von KI-Sicherheitsproblemen ist ihre Unsichtbarkeit. Klassische Angriffe generieren Signale: ungewöhnliche Netzwerkverbindungen, fehlerhafte Anfragen und/oder verdächtige Prozesse. Diese Signale lassen sich überwachen, korrelieren und alarmieren. KI-Fehlentscheidungen tun das nicht.

Ein System, das ein Objekt falsch klassifiziert, erzeugt keinen Fehler. Ein Modell, das durch manipulierte Trainingsdaten verzerrt wurde, verhält sich konsistent – nur eben falsch. Ein Sprachmodell, das vertrauliche Informationen ausgibt, folgt seiner internen Wahrscheinlichkeitslogik. Für Monitoring-Systeme ist das Normalbetrieb.

Diese Unsichtbarkeit führt dazu, dass viele KI-Sicherheitsprobleme erst dann auffallen, wenn der Schaden bereits eingetreten ist. In sicherheitskritischen Kontexten – Medizin, Mobilität, industrielle Steuerung – kann das fatale Folgen haben. Der Code ist korrekt, die Infrastruktur ist gehärtet, die Compliance erfüllt. Und dennoch trifft das System eine Entscheidung, die Menschen schadet.

1.5 Statistik ersetzt keine Sicherheit

Ein häufiges Missverständnis in der Diskussion um KI-Sicherheit ist die Gleichsetzung von hoher Genauigkeit mit Sicherheit. Modelle, die in Benchmarks gut abschneiden, gelten als zuverlässig. Fehlerraten werden statistisch analysiert, Konfidenzintervalle berechnet, Validierungsdaten ausgewertet. Doch diese Metriken sagen wenig über Sicherheit aus.

Ein Modell kann im Durchschnitt hervorragend performen und dennoch gezielt manipulierbar sein. Sicherheit bezieht sich nicht auf den Mittelwert, sondern auf Worst-Case-Szenarien. Genau dort versagen statistische Bewertungen. Adversariale Angriffe sind nicht zufällig. Sie sind gezielt konstruiert, um die Schwächen eines Modells auszunutzen. Ein

System, das in 99,9 % der Fälle korrekt arbeitet, kann in sicherheitskritischen Anwendungen unbrauchbar sein, wenn die verbleibenden 0,1 % systematisch provoziert werden können.

1.6 Compliance, Governance und Kontrollillusion

Mit der zunehmenden Regulierung von KI-Systemen wächst der Fokus auf Governance, Dokumentation und Prozesse. Modellkarten, Datenbeschreibungen, Risikoanalysen und Audit-Trails sollen Transparenz und Kontrolle schaffen. Diese Maßnahmen sind wichtig – aber sie lösen das Sicherheitsproblem nicht. Compliance beantwortet organisatorische Fragen: Wer ist verantwortlich? Welche Daten wurden verwendet? Welche Tests wurden durchgeführt? Sie beantwortet nicht die Frage, ob ein System gegen gezielte Angriffe resistent ist.

Ein vollständig dokumentiertes, auditierbares und regelkonformes KI-System kann dennoch hochgradig verwundbar sein. Governance beschreibt, wie ein System entwickelt wurde, nicht wie es sich unter Angriff verhält. Sicherheit erfordert technische Tests, kontinuierliche Beobachtung und explizite Bedrohungsmodelle.

1.7 Neue Angreifer, neue Denkweisen

Die Angreifer von KI-Systemen unterscheiden sich grundlegend von klassischen Hackern. Sie benötigen kein tiefes Verständnis von Betriebssystemen oder Netzwerken. Stattdessen analysieren sie Modelle,

Datenflüsse und Entscheidungslogiken. Sie experimentieren mit Eingaben, beobachten Reaktionen und leiten daraus Strategien ab.

Dieser Angreifer ist oft unsichtbar, skalierbar und automatisierbar. Angriffe lassen sich mit Skripts, Simulationen und generativen Modellen selbst durchführen. Die Grenze zwischen Nutzer und Angreifer verschwimmt. Sicherheitsstrategien, die auf Abschottung und Zugriffskontrolle setzen, greifen hier nicht.

1.8 Verschiebung der zentralen Sicherheitsfrage

Die zentrale Frage der klassischen IT-Sicherheit lautet: Kann ein Angreifer Zugriff auf mein System erlangen? Für KI-Systeme ist diese Frage zweitrangig. Die relevanten Fragen lauten:

- Kann ein Angreifer das Verhalten des Modells gezielt beeinflussen?
- Kann er Entscheidungen reproduzierbar fehlleiten?
- Kann er langfristige Effekte durch Datenmanipulation erzeugen?
- Kann er das System so nutzen, dass es seine eigenen Sicherheitsannahmen verletzt?

Diese Fragen lassen sich nicht mit Firewalls, Patches oder Penetrationstests beantworten. Sie erfordern ein Sicherheitsverständnis, das Modelle, Daten und Integration gemeinsam betrachtet.

1.9 Konsequenzen für Architektur und Betrieb

Wer KI-Systeme wie klassische Software absichert, erhält ein formal korrektes, aber praktisch verwundbares System. Sicherheit muss von Anfang an in Architektur, Datenmanagement und Betrieb integriert werden. Das betrifft nicht nur einzelne Komponenten, sondern den gesamten Lebenszyklus. KI-Sicherheit ist kein Feature, das man nachrüstet. Sie ist eine Eigenschaft des Gesamtsystems.

Um KI-Systeme wirksam abzusichern, muss man verstehen, wo sie angreifbar sind. Diese Angriffsflächen verteilen sich über den gesamten Lebenszyklus: von der Datensammlung über das Training bis zur Inferenz und Integration in produktive Systeme.

Im nächsten Kapitel betrachten wir diesen Lebenszyklus nicht aus Entwickler-, sondern aus Angreiferperspektive – und zwar als zusammenhängende Kette von Schwachstellen.

2 Der KI-Lebenszyklus als Angriffsfläche

Um die Risiken zu erfassen, die mit KI in der Praxis verknüpft sind, ist insbesondere die Betrachtung des KI-Lebenszyklus sinnvoll.

2.1 Sicherheit beginnt nicht im Code

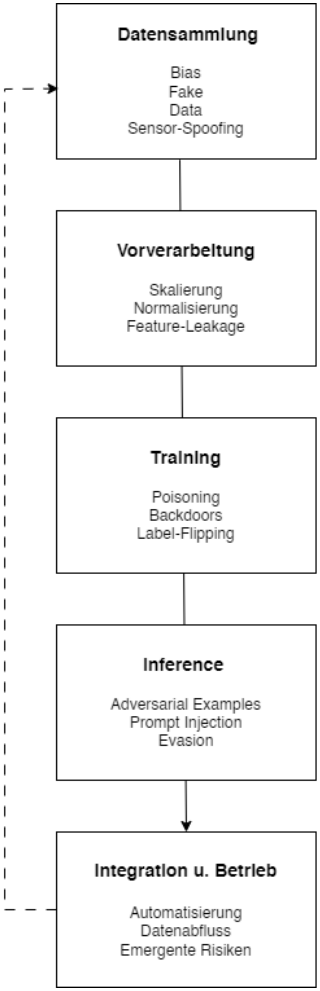
In klassischen Softwareprojekten beginnt Sicherheit dort, wo Code geschrieben wird. Schwachstellen entstehen durch fehlerhafte Implementierungen, unsichere Konfigurationen oder mangelhafte Zugriffskontrollen. Entsprechend konzentrieren sich Sicherheitsmaßnahmen auf Quellcode, Laufzeitumgebung und Netzwerkgrenzen. Dieses Modell setzt voraus, dass das Verhalten eines Systems weitgehend durch seinen Code determiniert ist.

Bei KI-Systemen ist diese Annahme nicht mehr haltbar.

Das Verhalten eines KI-Systems ergibt sich nicht allein aus Programmcode, sondern aus dem Zusammenspiel von Daten, Trainingsprozessen, Modellarchitektur, Parametern, Kontextinformationen und kontinuierlicher Interaktion mit der Umwelt. Sicherheit ist damit keine Eigenschaft einzelner Komponenten, sondern des gesamten Lebenszyklus. Jede Phase dieses Lebenszyklus beeinflusst das spätere Systemverhalten – und jede Phase eröffnet potenzielle Angriffsflächen.

Ein Angreifer, der KI-Systeme kompromittieren will, muss nicht auf den Produktivbetrieb warten. Oft ist es effektiver, deutlich früher

anzusetzen, dort wo Sicherheitsannahmen noch implizit sind oder gar nicht existieren.



Angriffsflächen bei KI-Systemen verteilen sich über den gesamten Lebenszyklus. Sicherheitsrelevante Manipulationen können bereits bei der Datensammlung entstehen und sich über Vorverarbeitung, Training, Inferenz und Systemintegration bis in den produktiven Betrieb fortsetzen.

Quellenverzeichnis

Barrett, C., Boyd, B., Bursztein, E., Carlini, N., Chen, B., Choi, J., ... & Yang, D. (2023). Identifying and mitigating the security risks of generative ai. Foundations and Trends® in Privacy and Security, 6(1), 1-52.

Bertino, E., Kantarcioglu, M., Akcora, C. G., Samtani, S., Mittal, S., & Gupta, M. (2021, April). AI for Security and Security for AI. In Proceedings of the eleventh ACM conference on data and application security and privacy (pp. 333-334).

Biggio, B., & Roli, F. (2018, October). Wild patterns: Ten years after the rise of adversarial machine learning. In Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security (pp. 2154-2156).

Carlini, N., Athalye, A., Papernot, N., Brendel, W., Rauber, J., Tsipras, D., ... & Kurakin, A. (2019). On evaluating adversarial robustness. arXiv preprint arXiv:1902.06705.

Chen, X., Liu, C., Li, B., Lu, K., & Song, D. (2017). Targeted backdoor attacks on deep learning systems using data poisoning. arXiv preprint arXiv:1712.05526.

Elliott, D., & Soifer, E. (2022). AI technologies, privacy, and security. Frontiers in Artificial Intelligence, 5, 826737.

Gambacorta, L., & Shreeti, V. (2025). The AI supply chain. BIS Papers.

Golda, A., Mekonen, K., Pandey, A., Singh, A., Hassija, V., Chamola, V., & Sikdar, B. (2024). Privacy and security concerns in generative AI: a comprehensive survey. IEEE Access, 12, 48126-48144.

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. arXiv preprint arXiv:1412.6572.

Greshake, K., Abdelnabi, S., Mishra, S., Endres, C., Holz, T., & Fritz, M. (2023, November). Not what you've signed up for: Compromising real-world llm-integrated applications with indirect prompt injection. In Proceedings of the 16th ACM workshop on artificial intelligence and security (pp. 79-90).

Habbal, A., Ali, M. K., & Abuzaraida, M. A. (2024). Artificial Intelligence Trust, risk and security management (AI trism): Frameworks, applications, challenges and future research directions. Expert Systems with Applications, 240, 122442.

Horowitz, M. C., Allen, G. C., Saravalle, E., Cho, A., Frederick, K., & Scharre, P. (2022). Artificial intelligence and international security. Center for a New American Security.

Hu, H., Salcic, Z., Sun, L., Dobbie, G., Yu, P. S., & Zhang, X. (2022). Membership inference attacks on machine learning: A survey. ACM Computing Surveys (CSUR), 54(11s), 1-37.

Hu, Y., Kuang, W., Qin, Z., Li, K., Zhang, J., Gao, Y., ... & Li, K. (2021). Artificial intelligence security: Threats and countermeasures. ACM Computing Surveys (CSUR), 55(1), 1-36.

Kalodanis, K., Rizomiliotis, P., & Anagnostopoulos, D. (2024). European artificial intelligence act: an AI security approach. Information & Computer Security, 32(3), 265-281.

Li, J. H. (2018). Cyber security meets artificial intelligence: a survey. Frontiers of Information Technology & Electronic Engineering, 19(12), 1462-1474.

Oseni, A., Moustafa, N., Janicke, H., Liu, P., Tari, Z., & Vasilakos, A. (2021). Security and privacy for artificial intelligence: Opportunities and challenges. arXiv preprint arXiv:2102.04661.

Pieters, W. (2011). Explanation and trust: what to tell the user in security and AI?. Ethics and information technology, 13(1), 53-64.

Rao, B., Zhang, J., Wu, D., Zhu, C., Sun, X., & Chen, B. (2024). Privacy inference attack and defense in centralized and federated learning: A comprehensive survey. IEEE Transactions on Artificial Intelligence.

Sarker, I. H., Furhad, M. H., & Nowrozy, R. (2021). Ai-driven cybersecurity: an overview, security intelligence modeling and research directions. SN Computer Science, 2(3), 173.

Szegedy, C., Zaremba, W., Sutskever, I., Bruna, J., Erhan, D., Goodfellow, I., & Fergus, R. (2013). Intriguing properties of neural networks. arXiv preprint arXiv:1312.6199.

Wang, B., Yao, Y., Shan, S., Li, H., Viswanath, B., Zheng, H., & Zhao, B. Y. (2019, May). Neural cleanse: Identifying and mitigating backdoor attacks in neural networks. In 2019 IEEE symposium on security and privacy (SP) (pp. 707-723). IEEE.

Yampolskiy, R. V. (Ed.). (2018). Artificial intelligence safety and security. CRC Press.

Yazmyradov, S. (2024). A Comprehensive Review of AI Security: Threats, Challenges, and Mitigation Strategies. The International Journal of Internet, Broadcasting and Communication, 16(4), 375-384.

Zou, A., Wang, Z., Carlini, N., Nasr, M., Kolter, J. Z., & Fredrikson, M. (2023). Universal and transferable adversarial attacks on aligned language models. arXiv preprint arXiv:2307.15043.

Stichwortverzeichnis

G

H

I

K

L

M

N

O

P

Q

S

T

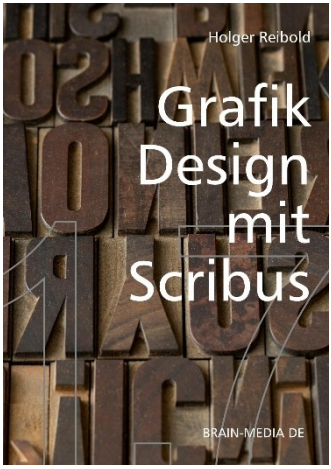
U

V

W

Z

Mehr von Brain-Media.de



Grafikdesign mit Scribus

In diesem Handbuch erfahren Sie alles, um mit Scribus ein professionelles Projekt umzusetzen – angefangen bei der Entwicklung kreativer Ideen bis zur konkreten Gestaltung.

Preis: 24,99 EUR

Umfang: 420 Seiten

Virtuelle Maschinen mit VirtualBox 7.x

So verwandeln Sie einen Rechner in ein ganzes Netzwerk oder bauen ein Testumgebung auf. Dieses Handbuch führt Sie in alle wichtigen Funktionen bis hin zur Cloud-Nutzung ein.

Preis: 16,99 EUR

Umfang: 150 Seiten

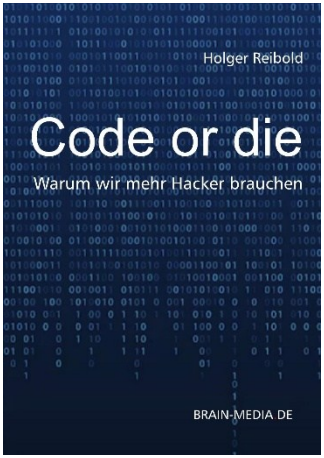
KI Red Teaming

Mit dem steigenden KI-Einsatz in Unternehmen wächst der Bedarf an Schutzmechanismen. Hier kommt das Red Team in Spiel.

Umfang: 160 Seiten

Preis: 19,99 EUR

Erscheint: Februar 2026



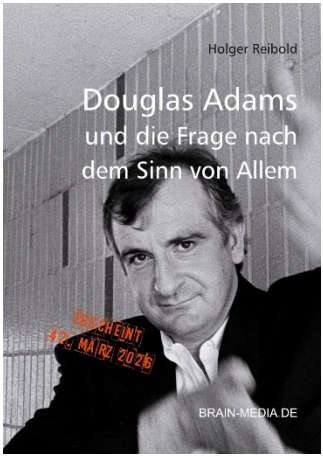
Code or die – Warum wir mehr Hacker brauchen

Ein Manifest für mehr digitale Selbstbestimmung, Neugierde und Eigenverantwortung. Medienkompetenzen alleine genügen nicht; die Gesellschaft von morgen braucht Digitalkompetenzen.

Umfang: 120 Seiten

Preis: 16,99 EUR

Erscheint Frühjahr 2026



42 – Douglas Adams und die Frage nach dem Sinn von Allem

Am 11. Mai 2026 ist Douglas Adams 25 Jahre tot. Der Kultautor hat der Welt wunderbar, skurrile Werke geschenkt. Jetzt ist es an der Zeit, den Autor kennenzulernen.

Umfang: 120 Seiten

Preis: 14,99 EUR

Erscheint: 42. März 2026

Towelday, das ultimative Handtuch für alle Fans

An seinem Todestag, dem Towelday, erinnern sich Fans an Douglas Adams und huldigen dem Kultautor.

100 % intergalaktisch geprüfte Baumwolle, nachhaltig Produktion zum Preis von 42 EUR.