

Vorbereitungshandbuch

Ausgabe 202611

Copyright © EXIN Holding B.V. 2026. All rights reserved.
EXIN® is a registered trademark of EXIN.

No part of this publication may be reproduced, stored, utilized or transmitted in any form or by any means, electronic, mechanical, or otherwise, without the prior written permission from EXIN.

Inhalt

1. Überblick 4
2. Prüfungsanforderungen 8
3. Liste der Grundbegriffe 12
4. Literatur 16
5. Karrierepfad 18

1. Überblick

EXIN AI Security Professional based on the OWASP AI Exchange (AISP.DE)

Anwendungsbereich

Die Zertifizierung EXIN AI Security Professional based on the OWASP AI Exchange bestätigt, dass die Fachkraft in der Lage ist, KI-Systeme während des Entwurfs, der Entwicklung, der Bereitstellung und des täglichen Einsatzes im Einklang mit den Richtlinien und Standards des Unternehmens sowie den geltenden Gesetzen sicher zu halten.

Diese Zertifizierung deckt folgende Themen ab:

- KI-Sicherheit innerhalb der Organisation
- KI-Sicherheitsbedrohungen
- KI-Sicherheitskontrollen
- KI-Sicherheitstests
- Privatsphäre und Compliance in der KI-Sicherheit

Zusammenfassung

Künstliche Intelligenz (KI) beeinflusst heutzutage Entscheidungen, Inhalte und Automatisierung in allen Branchen, was neue Möglichkeiten eröffnet, aber auch die Angriffsfläche vergrößert. Infolgedessen sehen sich Unternehmen mit sich schnell entwickelnden Risiken konfrontiert, darunter Angriffe, die die Manipulation von KI-Systemen, den Missbrauch von Modellen oder tool- und anbieterübergreifende Sicherheitslücken beinhalten, während Datenschutzanforderungen und gesetzliche Vorgaben immer strenger werden. In diesem Umfeld darf Sicherheit kein nachträglicher Einfall sein; sie muss in die Art und Weise integriert werden, wie KI entworfen, bereitgestellt und überwacht wird.

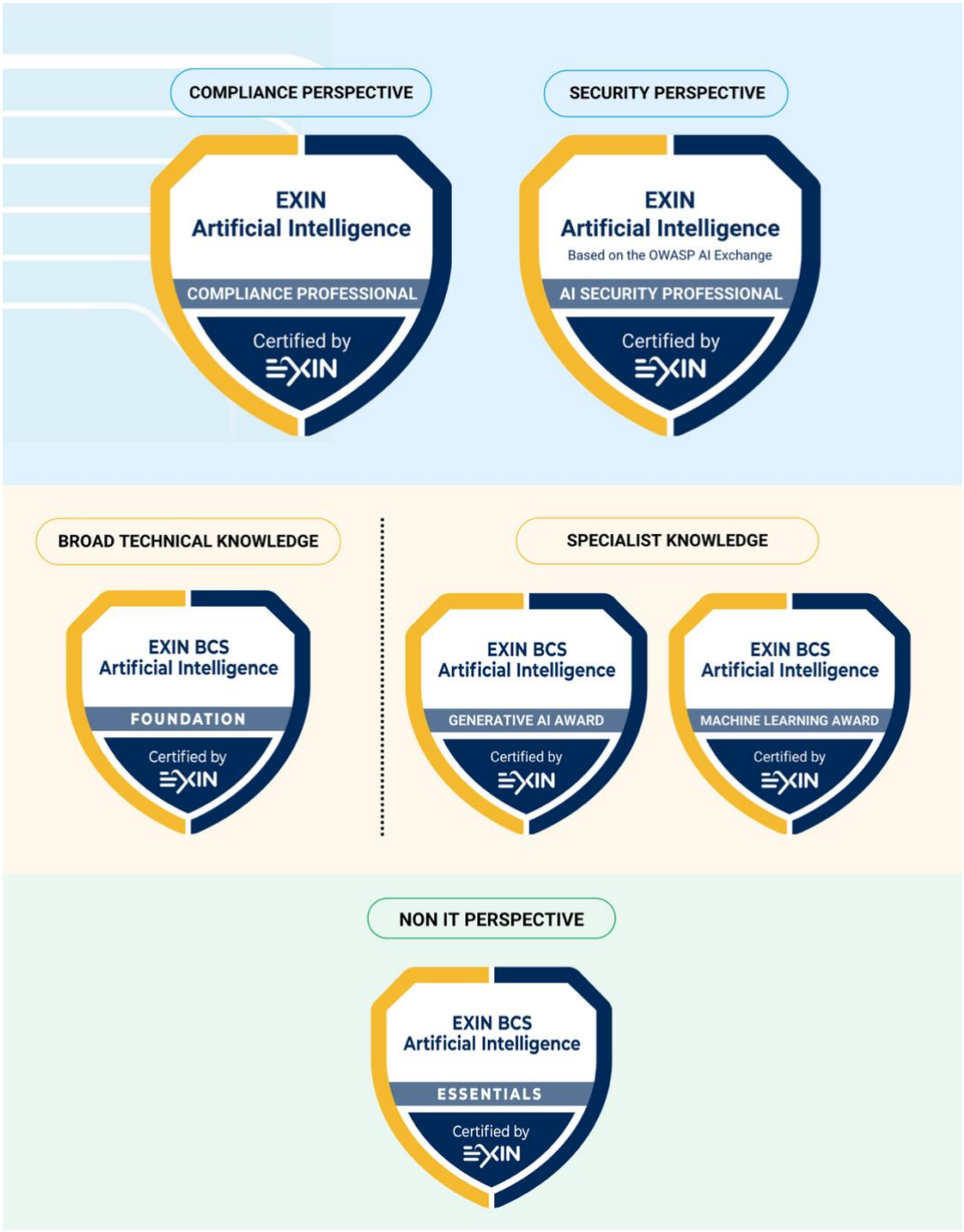
In den letzten Jahren hat sich KI-Sicherheit zu einer entscheidenden Priorität entwickelt. Die OWASP AI Exchange definiert Sicherheit als die Verhinderung von unbefugtem Zugriff, unbefugter Nutzung, Offenlegung, Störung, Veränderung oder Zerstörung, wobei „Veränderung“ auch die unerwünschte Manipulation des Verhaltens eines KI-Modells umfasst. Da KI zunehmend in Entscheidungsprozesse integriert wird, ist die Gewährleistung ihrer Zuverlässigkeit und Integrität unerlässlich, um Vertrauen zu schaffen und Risiken in diesem komplexen Sicherheitsumfeld zu minimieren.

Die auf der OWASP AI Exchange basierende Zertifizierung „EXIN AI Security Professional based on the OWASP AI Exchange“ bereitet Fachkräfte darauf vor, dieser Herausforderung zu begegnen. Sie stützt sich auf die OWASP AI Exchange, eine offene, gemeinschaftlich getragene Wissensdatenbank zu KI-Risiken, Bedrohungen, Kontrollen und Teststrategien. Die Zertifizierung bietet Fachkräften einen Weg, um KI-Sicherheit im Unternehmen, wichtige Bedrohungskategorien, wirksame Sicherheitskontrollen, KI-Sicherheitstests sowie Datenschutz und Compliance über den gesamten KI-Lebenszyklus hinweg besser zu verstehen.

Die AISP-Zertifizierung ist für Fachleute mit unterschiedlichem Hintergrund relevant und besonders attraktiv für diejenigen, die in Rollen wie Sicherheit und Privatsphäre, (KI-)Entwicklung, KI-Tests, Governance der KI-Sicherheit und Compliance tätig sind oder sich dafür interessieren. Inhaber dieser Zertifizierung verstehen, wie man KI-Sicherheitsprogramme entwirft, bewertet und verbessert, die widerstandsfähig, überprüfbar und datenschutzbewusst sind, und bringen dieses Wissen direkt in ihre tägliche Arbeit ein.

Kontext

Die Zertifizierung EXIN AI Security Professional ist Teil des EXIN Artificial Intelligence-Qualifizierungsprogramms.



Zielgruppe

Die Zertifizierung EXIN AI Security Professional based on the OWASP AI Exchange eignet sich für eine vielfältige Gruppe von Fachleuten, die mit KI-Systemen arbeiten, sowie für diejenigen, deren Schwerpunkt auf der Sicherheit und dem Datenschutz dieser Systeme liegt. Diese Rollen erfordern fundierte Kenntnisse, damit Organisationen vertrauenswürdige Anwendungen konzipieren, entwickeln, beschaffen, betreiben und warten können.

Dazu gehören unter anderem:

- KI-Fachleute
- (Cyber-)Sicherheitsexperten
- Datenschutz- und Privatsphäre-Experten
- Ingenieure für maschinelles Lernen (im Bereich Sicherheit)
- Dateningenieure
- Qualitäts- und Evaluierungsspezialisten
- Beauftragte für Governance, Risiko und Compliance

Zertifizierungsvoraussetzungen

- Erfolgreicher Abschluss der Prüfung EXIN AI Security Professional based on the OWASP AI Exchange

Eine Schulung wird dringend empfohlen. Kenntnisse der KI-Terminologie, zum Beispiel durch die Zertifizierung EXIN BCS Artificial Intelligence Essentials oder die Zertifizierung EXIN BCS Artificial Intelligence Foundation, werden dringend empfohlen.

Einzelheiten zur Prüfung

Art der Prüfung:	Multiple-Choice-Fragen
Anzahl der Fragen:	40
Mindestpunktzahl:	65 % (26/40 Fragen)
Einsicht in Dokumentation:	Nein
Notizen machen:	Nein
Elektronische Geräte/Hilfsmittel erlaubt:	Nein
Prüfungsdauer:	90 Minuten

Es gilt die Prüfungsordnung von EXIN.

Bloom Level

Die Zertifizierung EXIN AI Security Professional based on the OWASP AI Exchange testet Kandidatinnen und Kandidaten auf Bloom Level 2 und Level 3 nach der überarbeiteten Taxonomie von Bloom:

- Bloom Level 2: Verstehen – ein Schritt über das Wissen hinaus. Verstehen zeigt, dass Kandidatinnen und Kandidaten begreifen, was präsentiert wird, und bewerten können, wie der Unterrichtsstoff in ihrem eigenen Umfeld angewendet werden kann. Diese Art von Fragen soll zeigen, dass die Kandidatin oder der Kandidat in der Lage ist, die richtige Beschreibung von Fakten und Ideen zu organisieren, zu vergleichen, zu interpretieren und auszuwählen.
- Bloom Level 3: Anwenden – diese Stufe zeigt, dass die Kandidatin oder der Kandidat Inhalte in einem anderen als dem gelernten Kontext anwenden kann. Die Fragen zu dieser Lernstufe sollen zeigen, dass die Kandidatin oder der Kandidat Probleme in neuen Situationen lösen kann, indem sie oder er das erworbene Wissen bzw. die gelernten Tatsachen, Techniken und Regeln auf eine andere oder neue Art anwendet. Die Fragen beschreiben in der Regel ein kurzes Szenario.

Schulung

Präsenzstunden

Für diesen Kurs werden 14 Präsenzstunden empfohlen. Darin enthalten sind praktische Aufgabenstellungen oder Gruppenarbeiten, Prüfungsvorbereitung und kurze Pausen. Nicht enthalten sind: Mittagspausen, Hausaufgaben und die Prüfung.

Regelstudiendauer

84 Stunden (3 ECTS), je nach Vorwissen.

Akkreditierte Trainingsorganisationen

Eine Liste mit unseren akkreditierten Trainingsorganisationen finden Sie unter www.exin.com.

2. Prüfungsanforderungen

Die Prüfungsanforderungen sind im Einzelnen in den Prüfungsspezifikationen erläutert. In der unten dargestellten Tabelle finden Sie eine Liste mit den Themen (Prüfungsanforderungen) und Unterthemen (Prüfungsspezifikationen) des Moduls.

Prüfungs-anforderungen	Prüfungsspezifikationen	Gewichtung
1. KI-Sicherheit innerhalb der Organisation		**15 %**
	1.1 Organisation der KI-Sicherheit	10 %
	1.2 Bedrohungsmodellierung und Risiken durch agentische KI	5 %
2. KI-Sicherheitsbedrohungen		**37,5 %**
	2.1 Bedrohungen durch Eingaben	17,5 %
	2.2 Bedrohungen in der Entwicklungsphase	10 %
	2.3 Herkömmliche Sicherheitsbedrohungen während der Runtime	10 %
3. KI-Sicherheitskontrollen		**27,5 %**
	3.1 Governance	12,5 %
	3.2 Begrenzung sensibler Daten	7,5 %
	3.3 Begrenzung unerwünschten Verhaltens	7,5 %
4. KI-Sicherheitstests		**7,5 %**
	4.1 Bedrohungen im Testumfang	5 %
	4.2 Strategien für KI-Sicherheitstests	2,5 %
5. Privatsphäre und Compliance in der KI-Sicherheit		**12,5 %**
	5.1 Privatsphäre und KI-Sicherheit	5 %
	5.2 Compliance und Regulierung	7,5 %
		100 %

Prüfungsspezifikationen

1 KI-Sicherheit innerhalb der Organisation

1.1 Organisation der KI-Sicherheit

Die Kandidatin oder der Kandidat ist in der Lage…

1.1.1 die fünf G.U.A.R.D.-Schritte anzuwenden, um die KI-Sicherheit innerhalb einer Organisation zu organisieren.

1.1.2 zu verstehen, wie der verantwortungsvolle und vertrauenswürdige Einsatz von KI in der Organisation gefördert werden kann.

1.1.3 zwischen KI-Sicherheit und herkömmlicher (Cyber-)Sicherheit zu differenzieren.

1.1.4 KI-Assets und deren wichtigste Bedrohungen zu skizzieren.

1.2 Bedrohungsmodellierung und Risiken durch agentische KI

Die Kandidatin oder der Kandidat ist in der Lage…

1.2.1 die Schritte des Risikomanagements zur Durchführung einer Bedrohungsmodellierung mit Kontrollen für die KI-Sicherheit zu skizzieren.

1.2.2 typische Risiken im Zusammenhang mit agentischer KI zu erläutern.

2 KI-Sicherheitsbedrohungen

2.1 Bedrohungen durch Eingaben

Die Kandidatin oder der Kandidat ist in der Lage…

2.1.1 die verschiedenen Arten der Umgehung anhand des Wissensstands der Angreifer zu veranschaulichen: Zero-Knowledge Evasion (Umgehungsangriff mit Nullwissen), Perfect-Knowledge Evasion (Umgehungsangriff mit vollständigem Wissen), Partial-Knowledge Evasion (Umgehungsangriff mit Teilwissen), Transfer Attack (Transferangriff) und Evasion After Poisoning (Umgehung nach Vergiftung).

2.1.2 zwischen direkter und indirekter Prompt-Injection zu unterscheiden.

2.1.3 die sieben Schichten des Prompt-Injection-Schutzes zu veranschaulichen.

2.1.4 zu erläutern, wie es bei Modellausgaben, Model Inversion (Modellinversion) und Membership Inference zur Offenlegung sensibler Daten kommen kann.

2.1.5 zu erläutern, wie es zu einer Model Exfiltration (Modellausleitung) kommt.

2.1.6 Bedrohungen durch eingabegesteuerte Resource Exhaustion (Ressourcenerschöpfung) zu identifizieren.

2.2 Bedrohungen in der Entwicklungsphase

Die Kandidatin oder der Kandidat ist in der Lage…

2.2.1 Risiken durch Data Poisoning (Datenvergiftung) in der Entwicklungsphase zu identifizieren.

2.2.2 Direct Development-Time Model Poisoning (direkte Modellvergiftung in der Entwicklungsphase) zu erläutern.

2.2.3 das Supply-Chain Model Poisoning (Modellvergiftung in der Lieferkette) zu erläutern.

2.2.4 verschiedene Arten von Datenlecks in der Entwicklungsphase zu interpretieren: Development-Time Data Leak (Datenleck in der Entwicklungsphase), Direct Development-Time Model Leak (direktes Modellleck in der Entwicklungsphase) sowie Source Code Leak (Quellcode-Leck) und Configuration Leak (Konfigurationsleck).

2.3 Herkömmliche Sicherheitsbedrohungen während der Runtime
Die Kandidatin oder der Kandidat ist in der Lage…
2.3.1 zu erläutern, wie sich herkömmliche Sicherheitsbedrohungen während der Runtime auf KI-Komponenten auswirken können.
2.3.2 Direct Runtime Model Poisoning (direkte Modellvergiftung während der Runtime) und Direct Runtime Model Leak (direktes Modellleck während der Runtime) zu veranschaulichen.
2.3.3 die Sicherheitsbedrohungen zu beschreiben, die mit in KI-Ausgaben eingebetteten Injections und einem Input Data Leak (Leck von Eingabedaten) verbunden sind.
2.3.4 zu erläutern, inwiefern ein Leck von Augmentierungsdaten und eine Manipulation von Augmentierungsdaten Sicherheitsbedrohungen darstellen.

3 KI-Sicherheitskontrollen

3.1 Governance
Die Kandidatin oder der Kandidat ist in der Lage…
3.1.1 allgemeine Governance-Kontrollen für die Aufsicht über die KI-Sicherheit anzuwenden.
3.1.2 den Anwendungsbereich allgemeiner Governance-Kontrollen zu erläutern.
3.1.3 zu verstehen, wie sich die Implementierung der Kontrollen zwischen einem externen Anbieter eines KI-Modells und dem Entwickler oder Betreiber des KI-Systems (vorgefertigtes Modell) aufteilt.

3.2 Begrenzung sensibler Daten
Die Kandidatin oder der Kandidat ist in der Lage…
3.2.1 Kontrollen zur Begrenzung sensibler Daten zu implementieren, um die Vertraulichkeit und Integrität zu verbessern.
3.2.2 zu verstehen, wie die Begrenzung sensibler Daten Risiken verringert.

3.3 Begrenzung unerwünschten Verhaltens
Die Kandidatin oder der Kandidat ist in der Lage…
3.3.1 Kontrollen zu implementieren, um die Auswirkungen unerwünschten Modellverhaltens zu begrenzen.
3.3.2 zu verstehen, wie die Begrenzung unerwünschten Modellverhaltens die Leistung verbessert und Risiken verringert.

4 KI-Sicherheitstests

4.1 Bedrohungen im Testumfang
Die Kandidatin oder der Kandidat ist in der Lage…
4.1.1 zu erläutern, warum KI-Sicherheitstests wichtig sind und welchen Umfang sie haben.
4.1.2 zu verstehen, auf welche Bedrohungen über herkömmliche Sicherheitstests hinaus geprüft werden muss, wenn prädiktive oder generative KI zum Einsatz kommt.

4.2 Strategien für KI-Sicherheitstests
Die Kandidatin oder der Kandidat ist in der Lage…
4.2.1 Strategien für KI-Sicherheitstests unter Anwendung des allgemeinen Testansatzes zu beschreiben.

5 Privatsphäre und Compliance in der KI-Sicherheit

5.1 Privatsphäre und KI-Sicherheit

Die Kandidatin oder der Kandidat ist in der Lage…

5.1.1 Datenschutz und Bedenken hinsichtlich der Privatsphäre im Zusammenhang mit KI-Systemen zu erläutern.

5.1.2 Datenschutzgrundsätze für KI-Systeme auf ein gegebenes Szenario anzuwenden.

5.2 Compliance und Regulierung

Die Kandidatin oder der Kandidat ist in der Lage…

5.2.1 zu erläutern, wie die Normen ISO/IEC 23894, ISO/IEC 27005, ISO/IEC 42001 und ISO/IEC 5338 eine Organisation bei der Implementierung von Kontrollen und Prozessen unterstützen, die den geltenden KI-Vorschriften entsprechen.

5.2.2 die Herausforderungen im Zusammenhang mit der Compliance von KI-Systemen mit einschlägigen Gesetzen und Vorschriften, einschließlich der KI-Verordnung (KI-VO) und der DSGVO, zu verstehen.

5.2.3 Strategien zur Risikominderung bei Urheberrechtsverletzungen und zur Unterstützung der Compliance zu beschreiben.

3. Liste der Grundbegriffe

Diese Liste enthält Begriffe und Abkürzungen, mit denen die Kandidatinnen und Kandidaten vertraut sein sollten.

Bitte beachten Sie, dass die Kenntnis dieser Begriffe alleine nicht ausreicht. Die Kandidatin oder der Kandidat muss diese Begriffe auch verstehen und mit Beispielen belegen können.

Eine vollständige Liste der Grundbegriffe und ihrer Definitionen finden Sie im Glossar (Literatur B).

Englisch	**Deutsch**
accountability	Verantwortung
adversarial attack	adversarialer Angriff
adversarial testing	adversariales Testen
agentic AI	agentische KI
AI (artificial intelligence)	KI (künstliche Intelligenz)
AI agent	KI-Agent
AI assets	KI-Assets
AI jailbreak	KI-Jailbreak
AI management system (AIMS)	KI-Managementsystem (AIMS)
AI model	KI-Modell
AI model provider	Anbieter eines KI-Modells
AI privacy principles • accuracy • consent • data minimization and storage limitation • data subject rights • explainability • fairness and lawfulness • transparency • use limitation and purpose specification	Datenschutzgrundsätze für KI • Genauigkeit • Einwilligung • Datenminimierung und Speicherbegrenzung • Rechte der betroffenen Person • Erklärbarkeit • Fairness und Rechtmäßigkeit • Transparenz • Verwendungsbegrenzung und Zweckbestimmung
AI red teaming	KI-Red-Teaming
AI security	KI-Sicherheit
AI security governance	Governance der KI-Sicherheit
AI security testing approach	Ansatz für die KI-Sicherheitsprüfung
AI system	KI-System
AI system provider	Anbieter eines KI-Systems
anonymization	Anonymisierung
API (application programming interface)	API (Application Programming Interface)
audit	Audit
augmentation data	Augmentierungsdaten
automated oversight	automatisierte Aufsicht
automation	Automatisierung
bias	Bias
blast radius	Auswirkungsradius
chatbot	Chatbot
chief AI officer (CAIO)	Chief AI Officer (CAIO)
Chief Information Security Officer (CISO)	Chief Information Security Officer (CISO)
code	Code
compliance	Compliance

confidentiality	Vertraulichkeit
configuration leak	Konfigurationsleck
copyright	Urheberrecht
data leak	Datenleck
data protection	Datenschutz
data protection impact assessment (DPIA)	Datenschutz-Folgenabschätzung (DSFA)
data security	Datensicherheit
data subject	betroffene Person
denial-of-service (DoS)	Denial-of-Service (DoS)
denial-of-wallet (DoW)	Denial-of-Wallet (DoW)
deployer	Betreiber
development-time	Entwicklungsphase
development-time security threats • data poisoning during development-time • model poisoning during development-time • sensitive data leak during development-time	Development-Time Security Threats (Sicherheitsbedrohungen in der Entwicklungsphase) • Data Poisoning during Development-Time (Datenvergiftung in der Entwicklungsphase) • Model Poisoning during Development-Time (Modellvergiftung in der Entwicklungsphase) • Sensitive Data Leak during Development-Time (Leck sensibler Daten in der Entwicklungsphase)
embedding	Embedding
encryption	Verschlüsselung
ethical principles	Ethikgrundsätze
European Union (EU)	Europäische Union (EU)
G.U.A.R.D. (Govern, Understand, Adapt, Reduce, Demonstrate)	G.U.A.R.D. (Steuern, Verstehen, Anpassen, Reduzieren, Nachweisen)
General Data Protection Regulation (GDPR)	Datenschutz-Grundverordnung (DSGVO)
generative AI	generative KI
hijacking	Übernahme
human oversight	menschliche Aufsicht
inference-time	Inferenzzeit
information security management system (ISMS)	Informationssicherheits-Managementsystem (ISMS)
input data	Eingabedaten
input security threats • evasion • model exfiltration • prompt injection • resource exhaustion • sensitive data disclosure through use	Input Security Threats (Sicherheitsbedrohungen durch Eingabedaten) • Evasion (Umgehung) • Model Exfiltration (Modellausleitung) • Prompt-Injection (Prompt-Injektion) • Resource Exhaustion (Ressourcenerschöpfung) • Sensitive Data Disclosure through Use (Offenlegung sensibler Daten durch Nutzung)
integrity verification	Integritätsprüfung
intellectual property (IP)	geistiges Eigentum
ISO/IEC 42001	ISO/IEC 42001
large language model (LLM)	Large Language Model (LLM)
leakage	Leckage
machine learning	maschinelles Lernen
membership inference	Membership Inference

model architecture	Modellarchitektur
model inversion	Modellinversion
model leak	Modellleck
open source	quelloffen
output data	Ausgabedaten
output sanitization	Ausgabe-Sanitizing
output validation	Ausgabevalidierung
personal data	personenbezogene Daten
predictive AI	prädiktive KI
privacy	Privatsphäre
privacy by default	Datenschutz durch Voreinstellung
privacy by design	Datenschutz durch Technikgestaltung
privacy impact assessments (PIA)	Privacy Impact Assessments (PIA)
privacy incident	Datenschutzvorfall
processing (of personal data)	Verarbeitung (von personenbezogenen Daten)
prompt	Prompt
pseudonymization	Pseudonymisierung
publicly accessible	öffentlich zugänglich
ready-made model	vorgefertigtes Modell
residual risk	Restrisiko
responsible AI	verantwortungsvolle KI
retrieval-augmented generation (RAG)	Retrieval-Augmented Generation (RAG)
risk	Risiko
risk analysis	Risikoanalyse
risk levels for AI systems based on AI Act • minimal or no risk • limited risk • high risk • unacceptable risk	Risikostufen für KI-Systeme gemäß KI-VO • minimales oder kein Risiko • begrenztes Risiko • hohes Risiko • unannehmbares Risiko
risk management	Risikomanagement
risk mitigation	Risikominderung
risk rating	Risikoeinstufung
runtime	Runtime
runtime security threats • augmentation data manipulation • direct augmentation data leak • direct runtime model leak • direct runtime model poisoning • input data leak • output containing conventional injection	Runtime Security Threats (Sicherheitsbedrohungen während der Runtime) • Augmentation Data Manipulation (Manipulation von Augmentierungsdaten) • Direct Augmentation Data Leak (direktes Leck von Augmentierungsdaten) • Direct Runtime Model Leak (direktes Modellleck während der Runtime) • Direct Runtime Model Poisoning (direkte Modellvergiftung während der Runtime) • Input Data Leak (Leck von Eingabedaten) • Output containing Conventional Injection (Ausgabe mit klassischer Injection)
safety of an AI system	Safety eines KI-Systems
security breach	Sicherheitsverstoß

security controls • controls for sensitive data limitation • controls to limit the effect of unwanted behavior • governance controls	Sicherheitskontrollen • Kontrollen zur Begrenzung sensibler Daten • Kontrollen zur Begrenzung der Auswirkungen unerwünschten Verhaltens • Governance-Kontrollen
security incident	Sicherheitsvorfall
security risks	Sicherheitsrisiken
security testing	Sicherheitstests
security threats	Sicherheitsbedrohungen
(security) attack	(Sicherheits-)Angriff
sensitive data	sensible Daten
service provider	Service Provider
source code	Quellcode
source code leak	Quellcode-Leck
sponge attack	Sponge-Angriff
stakeholder	Stakeholder
startup	Start-up-Unternehmen
structured query language (SQL)	Structured Query Language (SQL)
tampering	Manipulation
third party	Dritter
threat modeling	Bedrohungsmodellierung
token	Token
tokenization	Tokenisierung
tool	Tool
training data	Trainingsdaten
trustworthy AI	vertrauenswürdige KI
unwanted behavior	unerwünschtes Verhalten
vulnerability	Sicherheitslücke

4. Literatur

Prüfungsliteratur

Das für die Prüfung benötigte Wissen wird durch folgende Literatur abgedeckt:

A. AI Security Community
OWASP AI Exchange
Open-Source-Publikation (April 2026)
Dieses Material ist frei verfügbar unter https://owaspai.org/.
Um die für die Zertifizierung verwendete PDF-Version (April 2026) herunterzuladen, gehen Sie zu www.exin.com. Klicken Sie auf ‚Professionals' und dann auf ‚Certifications', um die Zertifizierung zu finden. Den kostenlosen Download finden Sie unter ‚Required reading'.

Weiterführende Literatur

B. EXIN
Glossar
EXIN (2026)
Dieses Material ist frei verfügbar unter www.exin.com.

C. Europäische Union
Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 on harmonized rules on artificial intelligence (Artificial Intelligence Act).
Europäische Union (2024)
https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

D. ISO/IEC 23894:2023 (EN)
Information technology — Artificial intelligence — Guidance on risk management
Schweiz, ISO/IEC (2023)
https://www.iso.org/home.html

E. ISO/IEC 27005:2022 (EN)
Information security, cybersecurity and privacy protection — Guidance on managing information security risks
Schweiz, ISO/IEC (2022)
https://www.iso.org/home.html

F. ISO/IEC 42001:2023 (EN)
Information technology — Artificial intelligence — Management system
Schweiz, ISO/IEC (2023)
https://www.iso.org/home.html

G. ISO/IEC 5338:2023 (EN)
Information technology — Artificial intelligence — AI system life cycle processes
Schweiz, ISO/IEC (2023)
https://www.iso.org/home.html

Anmerkung

Weiterführende Literatur dient nur als Referenz und zur Vertiefung des Wissens.

Literaturmatrix

Prüfungs-anforderungen	Prüfungsspezifikationen	Literaturverweis
1. KI-Sicherheit innerhalb der Organisation		
	1.1 Organisation der KI-Sicherheit	A, Kapitel 0
	1.2 Bedrohungsmodellierung und Risiken durch agentische KI	A, Kapitel 0
2. KI-Sicherheitsbedrohungen		
	2.1 Eingabebedingte Bedrohungen	A, Kapitel 2
	2.2 Bedrohungen in der Entwicklungsphase	A, Kapitel 3
	2.3 Herkömmliche Sicherheitsbedrohungen während der Runtime	A, Kapitel 4
3. KI-Sicherheitskontrollen		
	3.1 Governance	A, Kapitel 0, 1
	3.2 Einschränkung sensibler Daten	A, Kapitel 1
	3.3 Eindämmung unerwünschten Verhaltens	A, Kapitel 1
4. KI-Sicherheitstests		
	4.1 Umfang der Bedrohungen	A, Kapitel 5
	4.2 Strategien für KI-Sicherheitstests	A, Kapitel 5
5. Privatsphäre und Compliance in der KI-Sicherheit		
	5.1 Privatsphäre und KI-Sicherheit	A, Kapitel 6
	5.2 Compliance und Regulierung	A, Kapitel 0, 1, 2, 3, 6

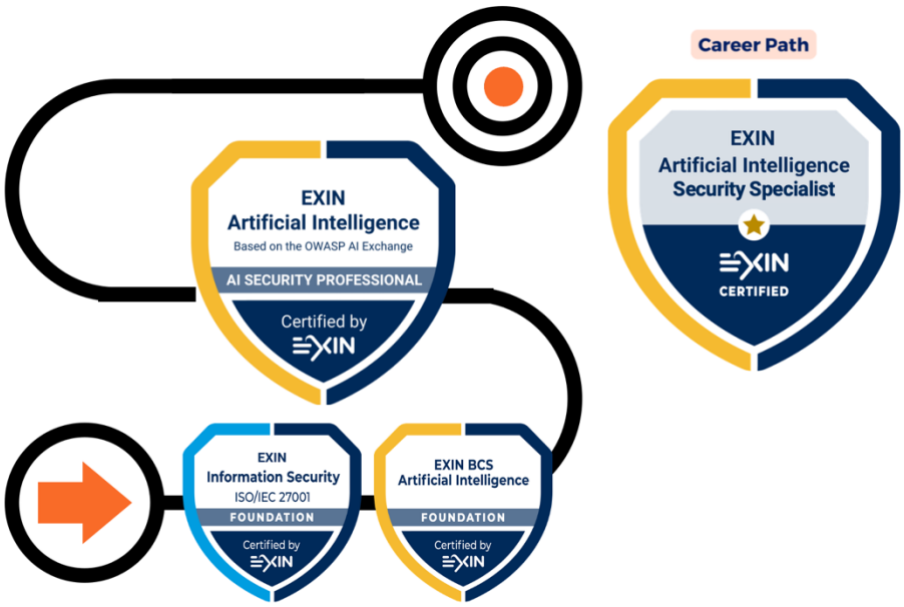
5. Karrierepfad

EXIN glaubt an den Wert des lebenslangen Lernens und an die Bedeutung, unterschiedliche Fähigkeiten zu kombinieren, um in der heutigen dynamischen und sich entwickelnden Welt erfolgreich zu sein. Mit unseren EXIN-Karrierepfaden können Kandidatinnen und Kandidaten sich auf bestimmte Jobrollen vorbereiten und ihre berufliche Entwicklung weiter vorantreiben. Weitere Informationen zu den EXIN-Karrierepfaden finden Sie auf unserer Website unter https://www.exin.com/career-paths/.

Die Zertifizierung EXIN AI Security Professional based on the OWASP AI Exchange ist Teil des folgenden EXIN-Karrierepfads.

EXIN AI Security Specialist

Der EXIN AI Security Specialist bestätigt, dass Fachkräfte sicher mit der Komplexität von KI-Technologien umgehen, die Datenintegrität wahren und globale Sicherheitsstandards einhalten können.



Certified for what's next

Kontakt EXIN

www.exin.com