

Practical Assignments

Edition 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.

Content

Introduction	4
Assignment 1: Organize AI security	6
Assignment 2: Attack techniques across the AI lifecycle	10
Evaluation	13
e-CF competences for EXIN AI Security Professional based on the OWASP AI Exchange	14

Introduction

Fulfillment of the practical assignments is not a mandatory part of the certification requirements for EXIN AI Security Professional based on the OWASP AI Exchange.

Both the trainer responsible for the practical assignments and the candidate must ensure that each candidate participates in the assignments in such a way that individual performance can be observed. This is particularly important in group assignments.

Adaptations

Accredited training organizations (ATOs) may set up practical assignments or simulations that fit the examples and use case studies presented in the training. The practical assignments do not need to follow the scenario's provided, although it is recommended. However, the scoring criteria may not be changed.

Adaptations can be assessed as part of the training materials audit on which the accreditation of the training organization is based. All assignments must be assessed based on the criteria provided by EXIN in the checklists in this document.

General Guidelines

In general, the following guidelines should be followed in addition to the requirements per assignment:

1. The candidate plans the activities within the assignment, unless otherwise stated.
2. The candidate communicates with others where needed.
3. The candidate provides their own solutions.
4. The candidate contributes to the quality of the assignment, especially in group work.
5. Solutions provided are realistic and thus fit for the scenario or case study.
6. Solutions provided match with business and IT objectives.

e-Competence Framework

The e-Competence Framework (e-CF)[1] is an accepted and common framework developed in Europe (www.ecompetences.eu). EXIN supports the e-CF, because EXIN believes in professionals showing their competences and growing towards their full potential by helping them make their competences transparent. The EXIN AI Security Professional based on the OWASP AI Exchange certification is based on the e-CF.

The practical assignments are used to demonstrate practical skills and experience which cannot be tested in a multiple-choice exam. Making the practical assignments part of the certification scheme helps to test the entire competence.

[1] The text is based on the European e-Competence Framework 3.0. CWA 16234:2014.

Assessment

Practical assignments can be assessed by an accredited trainer from the accredited training organization. It is the trainer's responsibility to familiarize themselves with the assessment criteria.

Each assignment has assessment criteria that are based on the exam specifications and linked to the e-CF. The criteria are found in the checklists. The trainer fills in the checklist for each individual candidate, stating whether or not the criteria have been observed.

The ATO should have evidence that practical assignments have been done in a particular training, this can be requested during an EXIN audit.

Assessment outside a training course

If your practical assignments cannot be assessed by a trainer, or are not assessed as part of your training, you can submit them directly to EXIN.

To have your assignments assessed:

- Copy the original headings of the assignments to your own working file.
- Complete all the practical assignments below their respective heading.
- Save the whole as a single PDF.
- Please do not use tables or pictures to present information.
- You can use lists to organize information.
- E-mail the PDF to candidatesupport@exin.com.
- Please use the following as the subject: **AISP – Practical Assignment assessment**

You will receive your result within 10 business days of EXIN receiving your assignments. EXIN will also send you a short report. The report includes feedback only on the parts that do not meet the assessment criteria.

Please note that you cannot keep re-submitting your practical assignments if you fail. You can re-submit your practical assignment once every 2 months if this is a pre-requirement for certification. Otherwise, your first submission result is final.

Assignment 1: Organize AI security

Background

YourLearning is a company that develops customized online learning applications for schools and training institutes. The company is adopting AI across all departments. You have joined YourLearning as an AI security officer.

An organization-wide review has identified the following AI use cases:

- **AI assistant:** This cross-departmental implementation should be able to draft e-mails and summarize meetings. It should also answer simple questions to support work.
- **Development:** This use of AI should support adding AI features to the flagship product of the company: an online AI trained to support learners in their individual journey.
- **Finance:** This AI implementation should be capable of processing invoices and expense claims, taking out as much manual work as possible.
- **HR:** The human resources (HR) department wants to use AI to screen job applications.
- **Marketing:** This team wants support with conducting market research and preparing weekly marketing e-mails.
- **Sales:** Sales wants to create customized presentations for clients to introduce YourLearning's products, and provide relevant updates and pricing information.

YourLearning has asked you to design an AI security approach and identify potential threats and controls to support the secure, responsible, and trustworthy use of AI.

Your assignment

Prepare a concise proposal to implement an AI security approach for YourLearning. Your proposal should be structured around the five G.U.A.R.D. steps and should address two of the AI use cases described above. Write it as a short report or slide set, focusing on clarity, practicality, and alignment with the organization's needs. You can follow these steps:

1. Build a structure
Apply the five G.U.A.R.D. steps to the organization and explain how they would work in practice for YourLearning. Briefly describe the actions required under each step, the main roles and responsibilities involved, and how AI use cases would be approved, monitored, and reviewed.

2. Identify AI assets and threats
Identify the most important AI-specific assets for the selected use cases and explain the main security threats associated with them. For each threat, assign an initial risk level of low, medium, or high.

Make sure your selection includes two of the use cases listed in the assignment, and that one of them is either the **AI assistant** or the **Development** use case. Your overview should include two assets for each use case and list one or two threats for each asset.

3. Threat modeling and risk treatment
Carry out a more detailed risk analysis for the two selected use cases. For each asset involved, answer the following questions:

- What could go wrong with this asset, security-wise?
- What could cause that problem?
- Who or what could be affected?
- What would happen if it did go wrong?
- How likely is it to happen?
- How serious would the impact be?
- What controls already exist?
- What additional controls could reduce the risk?

4. Define processes and controls
Define a concise set of organization-wide processes and security controls that YourLearning should implement across all AI use cases. These could cover:

- Access to AI systems
- Prevention of unnecessary sensitive-data input
- Logging, monitoring, and incident reporting
- Human oversight
- Employee training

5. AI security testing strategy
Describe how YourLearning should test its AI systems before deployment and during operation. Apply the general testing approach and include:

- What will be tested and why
- Tests for misuse, malicious input, sensitive-data exposure, and resource exhaustion
- The expected results and criteria for passing the tests
- How findings will be reported and resolved
- When the AI systems should be tested again

Recommended time

1 hour

Expected results

A concise report or slide set of approximately 5-8 slides or 3-4 pages, containing:

- 1 summary of the five G.U.A.R.D. steps, containing:
  - A clear description of what each step looks like for YourLearning.
- 1 asset-and-threat table covering 2 of the mentioned use cases, containing:
  - 2 assets within YourLearning
  - 1-2 threats per identified asset
  - One of the choices must include the AI Assistant or the Development use case.
- 1 risk analysis for the 2 selected use cases.
- 1 list of 3 controls for the selected use cases
- 1 short draft for an AI security testing approach

Assessment criteria for the exercise

The candidate should briefly justify the proposed risk levels and controls. Solutions must be realistic and support YourLearning's business and IT objectives.

In a group assignment, each candidate must explain an identifiable part of the work so that individual performance can be observed.

Assignment 1 Checklist

The trainer can assess each candidate on each exam specification.

e-Competence	Exam specification	The candidate can...	Observed? (Tick box if observed)
E.3., E.8.	1.1.1	apply the five G.U.A.R.D. steps to organize AI security within an organization.	☐
E.7.	1.1.4	outline AI-specific assets and their key threats.	☐
E.7.	1.2.1	outline risk management steps to conduct threat modeling with controls for AI security.	☐
E.7.	2.1.6	identify input driven resource exhaustion threats.	☐
E.8.	3.1.1	apply general governance controls for AI security oversight.	☐
A.5., E.8.	3.2.1	implement controls to limit sensitive data in order to enhance confidentiality and integrity.	☐
B.3., C.4.	4.2.1	describe AI security testing strategies by using the general testing approach.	☐

Assignment 2: Attack techniques across the AI lifecycle

Background

EasyGroceries is an online grocery store offering next-day delivery. Due to rapid growth, its customer service team is handling an increasing number of questions and complaints.

To reduce this workload, EasyGroceries plans to obtain a third-party AI chatbot. The chatbot will answer questions about deliveries, product availability, discount coupons, order status, and selected refunds. It will process customer messages and information retrieved from internal and external sources.

The solution uses third-party models, open-source components, external datasets, and automated development and deployment pipelines. It can receive model and configuration updates and connect to EasyGroceries' product catalogue, order-management platform, customer database, and refund-processing system.

Before deploying the chatbot, EasyGroceries wants to understand the security threats that could arise across its lifecycle. EasyGroceries' customer-service manager is also concerned that malicious instructions in customer messages or retrieved content could lead the chatbot to disclose order information or issue unauthorized refunds, and would like an explanation of how the seven layers of prompt injection can work together to reduce that risk.

Your assignment

Using the OWASP AI Security Threat Model One-Pager[2], analyze the EasyGroceries chatbot from an attacker's perspective. Prepare a report or slide set that identifies theoretically applicable threats and determines which risks EasyGroceries should treat before deployment. You can follow these steps:

1. Describe the system and its lifecycle
Provide a simple overview of the chatbot and its main components. Include:

- Inputs and outputs
- Data and models
- Third-party components
- Development and deployment processes
- Connected business systems
- Possible access points

2. Identify theoretically applicable threats
Identify threats that could theoretically affect the chatbot during development or operation. For each identified threat, provide a short example showing how it could apply to EasyGroceries. Each example should briefly identify the attacker or source, the attack method, the target, and the possible consequence.

[2] Full website: https://owaspai.org/images/threatmodelonepager.png

3. Assess and prioritize credible risks
Select 2 threats that represent the most credible risks to EasyGroceries.
For each selected threat:

- explain how the attack could occur
- assess its likelihood and impact as low, medium, or high
- explain why it should or should not be treated before deployment
- identify the assets, systems, or people that could be affected

4. Recommend controls
Recommend controls for each of the 2 selected risks. Include:

- At least 1 preventive control
- At least 1 control that detects or limits the impact of the attack

Briefly explain how EasyGroceries could check whether the proposed controls work as intended.

Recommended time

1 hour

Expected results

A report or slide set of 4-6 slides or 2-3 pages, containing:

- 1 brief overview of the chatbot and its main components
- 1 list of 4 theoretically applicable threats
- 1 short attack example for 2 selected threats.
- 1 prioritized risk estimation covering the 2 selected threats.
- 1 comparison of direct and indirect prompt injection, with 1 example of each
- 1 overview of the seven layers of prompt injection protection.
- At least 2 controls for each of the 2 selected risks:
  - 1 preventive control
  - 1 detective or impact-limiting control
- 1 brief justification of the likelihood, impact, prioritization, and treatment decisions

Assessment criteria for the exercise

Solutions must be realistic and appropriate for EasyGroceries' chatbot and business environment.

In a group assignment, each candidate must explain at least one attack technique, its risk assessment, and its proposed controls so that individual performance can be observed.

Assignment 2 Checklist

The trainer can assess each candidate on each exam specification.

e-Competence	Exam specification	The candidate can...	Observed? (Tick box if observed)
E.7.	2.1.1	illustrate the various types of evasion based on the attackers' access to knowledge: zero-knowledge evasion, perfect-knowledge evasion, partial-knowledge evasion, transfer attack, and evasion after poisoning.	☐
E.7.	2.1.2	distinguish between direct prompt injection and indirect prompt injection.	☐
A.5.	2.1.3	illustrate the seven layers of prompt injection protection.	☐
E.7.	2.2.1	identify data poisoning risks during development-time.	☐
E.7., E.8.	2.2.4	interpret different sensitive data leaks during development time: development time data leak, direct development time model leak, and source code/configuration leak.	☐
E.7.	2.3.2	illustrate direct runtime model poisoning and direct runtime model leak.	☐

Evaluation

The trainer can fill out the final evaluation below for each individual candidate. When a minimum of 8 out of 13 (62%) of the criteria have been observed, the candidate has successfully performed the practical assignments.

Please note that some of the exam specifications are assessed in more than one assignment. If the requirement has been observed in at least one of the assignments, the trainer may assess the exam specification as 'observed'.

e-Competence	Exam specification	The candidate can...	Observed? (Tick box if observed)
E.3., E.8.	1.1.1	apply the five G.U.A.R.D steps to organize AI security within an organization.	☐
E.7.	1.1.4	outline AI-specific assets and their key threats.	☐
E.7.	1.2.1	outline risk management steps to conduct threat modeling with controls for AI security.	☐
E.7.	2.1.1	illustrate the various types of evasion based on the attackers' access to knowledge: zero-knowledge evasion, perfect-knowledge evasion, partial-knowledge evasion, transfer attack, and evasion after poisoning.	☐
E.7.	2.1.2	distinguish between direct prompt injection and indirect prompt injection.	☐
A.5.	2.1.3	illustrate the seven layers of prompt injection protection.	☐
E.7.	2.1.6	identify input driven resource exhaustion threats.	☐
E.7.	2.2.1	identify data poisoning risks during development-time.	☐
E.7., E.8.	2.2.4	interpret different sensitive data leaks development-time: development time data leak, direct development time model leak, and source code/configuration leak.	☐
E.7.	2.3.2	illustrate direct runtime model poisoning and direct runtime model leak.	☐
E.8.	3.1.1	apply general governance controls for AI security oversight.	☐
A.5., E.8.	3.2.1	implement controls to limit sensitive data in order to enhance confidentiality and integrity.	☐
B.3., C.4.	4.2.1	describe AI security testing strategies by using the general testing approach.	☐
		Total	

Total number of specifications observed:	
Candidate passed / failed:	

e-CF competences for EXIN AI Security Professional based on the OWASP AI Exchange

You can find all of the e-Competence Framework competences related to the EXIN AI Security Professional based on the OWASP AI Exchange certification below. Also indicated is the level of the competence and whether the competence is covered entirely, partially or superficially. For more information about the e-CF, please visit www.ecompetences.eu or contact EXIN.



e-Competence Level		1	2	3	4	5
A.5.	Architecture Design					
B.3.	Testing					
C.4.	Problem Management					
E.3.	Information Security Startegy Development					
E.7.	Risk Management					
E.8.	Information Security Management					

Copyright © EXIN Holding B.V

Certified for what's next

Contact EXIN

www.exin.com