

Contents	Page
European foreword	5
Introduction	6
1 Scope	8
2 Normative references	8
3 Terms and definitions	8
3.1 Terms from EU AI Act and other EU regulations	8
3.2 Other related terms	10
4 Abbreviated terms	17
5 Cybersecurity framework for the AI system	17
5.1 General requirements	17
5.2 Relationship between threats, vulnerabilities and measures	18
5.3 Required outcomes	18
5.4 Documentation of the cybersecurity framework	18
6 Determination of relevant circumstances for AI cybersecurity	19
6.1 General requirements	19
6.2 Scope and operational context	19
6.3 Cybersecurity-relevant assets	19
6.4 Operating and deployment circumstances	20
7 Identification of AI-specific cybersecurity vulnerability	20
7.1 General requirements	20
7.2 Vulnerability identification process	20
7.3 Scope of AI-specific vulnerabilities	21
7.4 Link to threats and measures	21
8 Identification of relevant cybersecurity threats	21
8.1 General requirements	21
8.2 Threat identification process	21
8.3 Scope of AI-specific threats	22
8.4 Link to vulnerabilities and measures	23
9 Determination of relevant risks for AI cybersecurity	23
9.1 General requirements	23
9.2 Determination of cybersecurity risks relevance	24
9.3 Risk acceptance and linkage to AI system risk control measures	24
10 Select and implement measures	25
10.1 Data poisoning	25
10.1.1 Prevent	25
10.1.2 Detect	25
10.1.3 Respond	26
10.1.4 Resolve	26
10.1.5 Control	26
10.2 Model poisoning	27
10.2.1 Prevent	27
10.2.2 Detect	28
10.2.3 Respond	28
10.2.4 Resolve	28

10.2.5 Control	29
10.3 Adversarial attacks or model evasion	29
10.3.1 Prevent	29
10.3.2 Detect	30
10.3.3 Respond	31
10.3.4 Resolve	31
10.3.5 Control	31
10.4 Confidentiality attacks	32
10.4.1 Prevent	32
10.4.2 Detect	33
10.4.3 Respond	33
10.4.4 Resolve	33
10.4.5 Control	34
10.5 Model flaws	34
10.5.1 Prevent	34
10.5.2 Detect	35
10.5.3 Respond	35
10.5.4 Resolve	35
10.5.5 Control	36
10.6 Threats related to generative AI models (including LLMs)	36
11 Verification and testing requirements	36
11.1 General requirements	36
11.2 Life cycle triggers	37
11.3 Types of cybersecurity testing	37
11.4 Poisoning resistance testing	37
11.5 Adversarial testing	38
11.5.1 General requirements	38
11.5.2 Identification of relevant adversarial attacks	38
11.5.3 Test design and execution	38
11.5.4 Attacker knowledge assumptions	39
11.6 Confidentiality testing	39
11.7 Model or algorithm exploitation testing	39
11.8 Composite and red-team testing	40
11.9 Acceptance criteria	40
12 Documentation requirements for AI cybersecurity	40
12.1 General requirements	40
12.2 Documentation of relevant circumstances	40
12.3 Documentation of vulnerabilities	40
12.4 Documentation of threats	41
12.5 Documentation of cybersecurity risks	41
12.6 Documentation of technical measures and testing	41
12.7 Documentation of instructions for use	42
Annex A (informative) Explanatory guidance supporting Clauses 5 to 12	43
A.1 Purpose and use	43
A.2 Overview of AI cybersecurity across the life cycle	43
A.3 Relationship between clauses in this document	43
A.4 Illustration of AI cybersecurity activities	44
A.5 Application of AI cybersecurity measures	44
A.6 Interaction with other standards	45
A.7 Continuous improvement and evolution of threats	45
A.8 Examples of circumstances, context and environment	45
A.9 Examples of AI-specific threats	46
A.10 Examples of AI-specific vulnerabilities	47
A.11 Examples of measures	47
Annex B (informative) Recommended organizational controls	48
B.1 AI management system	48
B.2 Apply software engineering best practice processes in AI development	48
B.3 Include AI considerations in the information security management system	48

Annex C (informative) Other related standards	49
Annex D (normative) AI-specific assets	50
Annex ZA (informative) Relationship between this European Standard and the essential requirements of Regulation 2024/1689 aimed to be covered	52
Bibliography	55