

Inhaltsverzeichnis

Vorwort der Projektleiter	1
Was ist neu in den Top 10 von 2025	1
Die Zukunft	2
Das deutsche Übersetzungsteam	3
Über diese Übersetzung	3
LLM01:2025 Prompt Injection	4
Beschreibung	4
Arten von Prompt Injection- Schwachstellen	5
Präventions- und Mitigationsstrategien	6
Beispiele für Angriffsszenarien	8
Referenzlinks	10
Verwandte Frameworks und Taxonomien	11
LLM02:2025 Offenlegung sensibler Informationen	12
Beschreibung	12
Gängige Beispiele für Schwachstellen	13
Präventions- und Mitigationsstrategien	13
Beispiele für Angriffsszenarien	16
Referenzlinks	17
Verwandte Frameworks und Taxonomien	17
LLM03:2025 Lieferkette	18
Beschreibung	18
Gängige Beispiele für Risiken	19
Präventions- und Mitigationsstrategien	22
Beispiele für Angriffsszenarien	24
Referenzlinks	28
Verwandte Frameworks und Taxonomien	28
LLM04: Poisoning von Daten und Modellen	29
Beschreibung	29

Gängige Beispiele für Schwachstellen	30
Präventions- und Mitigationsstrategien	30
Beispiele für Angriffsszenarien	31
Referenzlinks	32
Verwandte Frameworks und Taxonomien	33
LLM05:2025 Fehlerhafte Ausgabeverarbeitung	34
Beschreibung	34
Gängige Beispiele für Schwachstellen	35
Präventions- und Mitigationsstrategien	35
Beispiele für Angriffsszenarien	36
Referenzlinks	38
LLM06:2025 Übermäßige Handlungsfreiheit	39
Beschreibung	39
Gängige Beispiele für Risiken	40
Präventions- und Mitigationsstrategien	41
Beispiele für Angriffsszenarien	45
Referenzlinks	45
LLM07:2025 Offenlegung des Systems Prompts	46
Beschreibung	46
Gängige Beispiele für Risiken	47
Präventions- und Mitigationsstrategien	49
Beispiele für Angriffsszenarien	50
Referenzlinks	50
Verwandte Frameworks und Taxonomien	51
LLM08:2025 Schwachstellen in Vektoren und Embeddings	52
Beschreibung	52
Gängige Beispiele für Risiken	52
Kontextübergreifende Informationslecks und Wissenskonflikte in der Föderation	54
Präventions- und Mitigationsstrategien	54
Beispiele für Angriffsszenarien	55
Referenzlinks	57
LLM09:2025 Fehlinformationen	58
Beschreibung	58
Gängige Beispiele für Risiken	59
Präventions- und Mitigationsstrategien	60
Beispiele für Angriffsszenarien	62

Referenzlinks	62
Verwandte Frameworks und Taxonomien	63
LLM10:2025 Unbegrenzter Verbrauch	64
Beschreibung	64
Gängige Beispiele für Schwachstellen	64
Präventions- und Mitigationsstrategien	66
Beispiele für Angriffsszenarien	69
Referenzlinks	70
Verwandte Frameworks und Taxonomien	71