

ChatGPT GPT-4 Claude DeepSeek

(Large Language

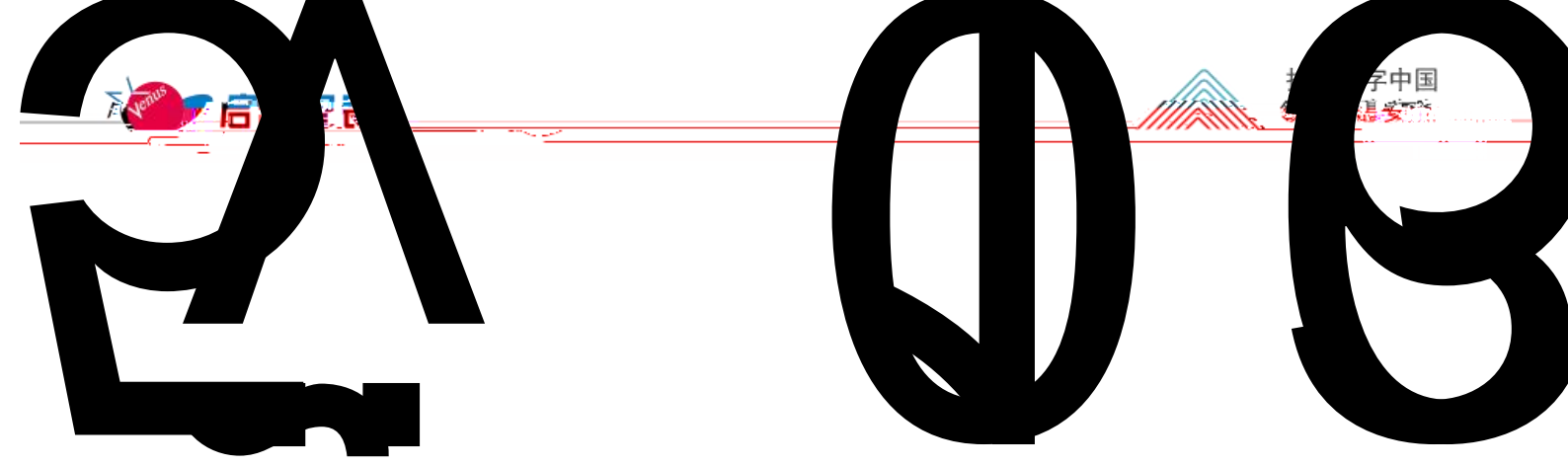
Model s, LLMs)

AI

"

"

n



j j̈ Y Ÿ 5

L



TensorFlow PyTorch Numpy



API

API

QIama

"

" " "

APP

DeepSeek APP

APP

2.2

●

“ ”

2.3

2.4

AI

AI

AI

Web

(

XSS)

AI

(

AI

)

AI

AI

AI

AI

- Web

Web

API

XSS

SQL

CSRF

API

- API

●

AI

AI - R-IAM

AI AI Identity

Non-Human Identity, NHI

RAG

IAM

RAG

API

API

()

I AM

- 1 AI 安全与隐私保护 AI - R-IAM . 2025 2 .
- 2 AI 安全与隐私保护 Demon AI Security Handbook 2025 5 .
- 3 DeepSeek 安全与隐私保护 . 2025 4 .
- 4 OWASP Top 10: LLM & Generative AI Security Risks <https://genai.owasp.org>.

EU AI ACT

2024

5 8 2025 2

AI AI

AI AI

- ISO/IEC 42001 ()

ISO IEC AI

AI MS

AI AI

- IEEE AI (IEEE AI Ethics Guidelines)

IEEE AI

- /

" "

" "

2023 8

1.0 2024 9

4

2

AI

- OWASP Top 10 for LLM Applications 2025

Web

Qu "

- MIT AI Risk Repository

Google DeepMind

FSF Frontier Safety Framework

DeepMind

AI

SAI F

FSF

● OpenAI AI

OpenAI LLM

AI

"Preparedness Framework"

AI