

- Analysis
  - AI
    - AI
    - ML
    - DL/NN
    - NLP
    - LLM
    - GPT
  - prompt engineering
    - few shot prompting
    - chain of thought prompting
  - jailbreaking
    - Prompt injection
    - Prompt leaking
    - DAN (Do Anything Now)
    - Roleplay
    - Developer mode
  - risks
    - ethical
      - misinformation
      - identity theft
      - bias amplification
      - copyright violation
      - military use
    - moral
      - generation of harmful/explicit content
    - cybersecurity
      - malware generation
      - ai-powered malware
      - phishing & deepfake
  - content moderation
    - prompt level
      - prompt filtering
      - prompt transformation
      -
    - model level
      - safety fine-tuning
      - pruning
      - moving target defense
  - legislation
    - EU
      - AI Act
      - Minimal risk
      - Limited risk
      - High risk
      - Unacceptable risk

- US
  - state level legislation
    - example: California
  - no federal law
- China
  - 3 regulations
  - Algorithm recommendation
  - Deep synthesis
  - Generative AI