AI Threat Map

Author: Sandy Dunn
sandy@quarkiq.com v1.6 212024

Threat Using AI Models

- LLM01: Prompt Injection
- LLM02: Insecure Output Handing
- LLM03: Trained Data Poisoning
- LLM05: Supply Chain Attack
- LLM06: Sensitive Information Disclosure
- LLM07: Insecure Plugin Design
- LLM08: Excessive Agency
- LLM09: Overreliance
- Indirect Prompt Injection
  - Passive
  - Active
  - User-driven Injection
  - Hidden injection
  - Payload Splitting
- Fake Resources
- Copyright infringement

Threats to AI Models

- LLM04: Denial of Service
- LLM10: Model Theft
- ML03: Model Inversion Attack
- ML07: Transfer Learning Attack
- ML08: Model Skewing Attack
- ML10: Model Poisoning
- Inadequate AI Alignment
- Improper Error Handling
- Robust multi-prompt and multi-model attacks
- Traditional Attacks

Threats from AI Models

- Misidentification i.e. wrongful arrest
- False Information i.e. criminal offenses
- Misinformation influence i.e. elections
- Private Information used in training
- Deep Fakes
  - Disinformation campaigns
  - Abused authentication
  - Synthetic or composite fakes
- Shallow Fake
  - Slightly altered fake image
- Attack Acceleration
  - FraudGPT
  - DarkBARD
  - DarkGPT
  - PoisonGPT
  - DarkBERT
  - DAN 9.0
  - ChaosGPT
- Hallucination Squatting
- Artificial Consciousness
- Honey or Poisoned Characters
- Social manipulation

Threats NOT using AI Models

- Competitive Disadvantage
- Limited Customer Engagement: Inablility to scale personalized communication
- Innovation Stagnation: Slower pace of innovation and improvements
- Operational Inefficiency: Slower and less efficient processes
- Market Perception: Viewed as outdated by customers and partners
- Higher risk of human error in processes
- Inefficient allocation of human resources

AI Legal & Regulatory Threats

- Legal
  - Product Warranties
  - Indemnification
  - Copyright
  - Licensing
  - Legal Obligations (e.g. fiduciary responsibility)
  - Privacy
- Regulatory
  - US Federal
    - Department of Justice (DOJ)
    - Consumer Financial Protection Bureau (CFPB)
    - Federal Trade Commission (FTC)
    - Equal Employment Opportunity Commission (EEOC)
  - Against AI Profiling
    - California
    - Colorado
    - Connecticut
    - Maryland (hiring only)
    - New Jersey (civil rights, employment)
    - New York (hiring only)
    - Virginia
    - Tennessee
  - US State Law
    - US State Law chatbot notification
      - California
    - US State Law Privacy
      - Colorado
      - Connecticut
      - Indiana
      - Iowa
      - Montana
      - Oregon
      - Tennesse
      - Texas
      - Utah
      - Virginia
      - Washington
    - US State Law Biometrics
      - Florida
      - Illinois
      - Washington
      - Maryland
      - Vermont
  - EU AI Act
  - Canada GenAI Guardrails
  - China GenAI Measures
  - Peru Law six core principles
  - Spain AESIA
  - South Korea Digital Bill of Rights