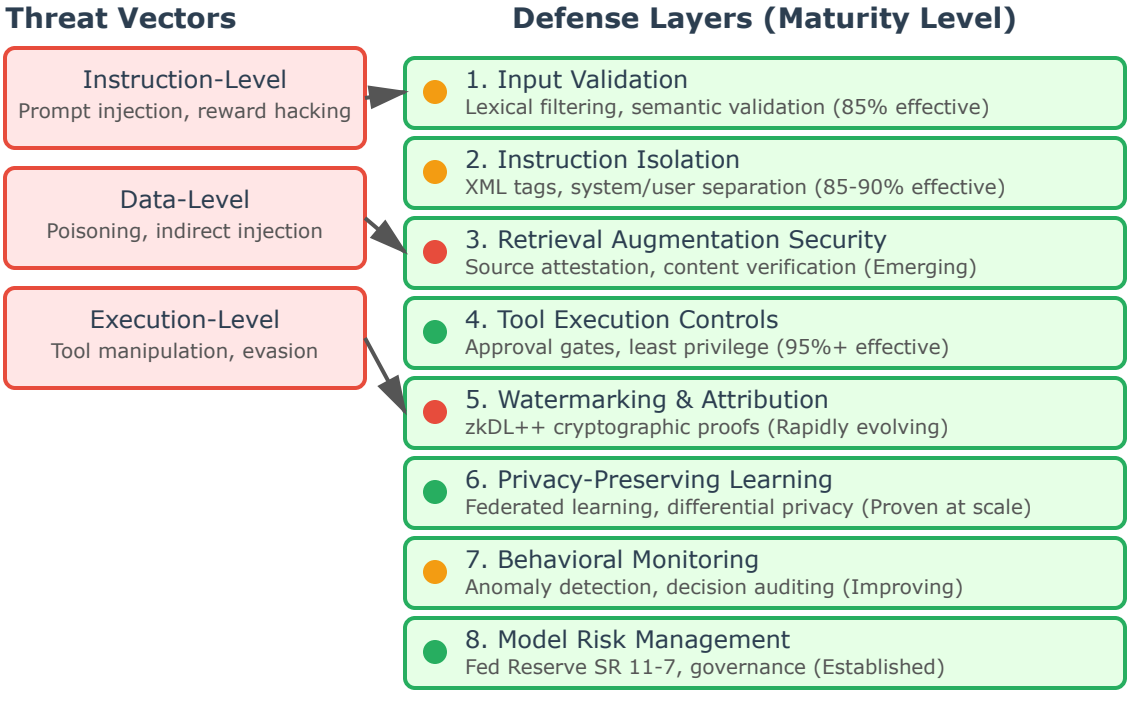
Defense-in-Depth Security Architecture

Eight complementary layers protecting autonomous systems



2026 Security Landscape

Threat Evolution:

- 58% of attacks in 2026 will be agentic-driven (first time majority non-human)
- OWASP Top 10 LLM 2025: Prompt injection remains #1 critical vulnerability
- $10B+ annual AI model IP theft; 98% fidelity model extraction possible

Defense Maturity:

- Tool execution controls: 95%+ effective (most mature layer)
- RAG security & watermarking: Rapidly evolving (monthly breakthroughs)

Maturity:

- Established
- Moderate
- Emerging