TruAID

Untrusted Agents In The Wild

Translator

Translate ...

Writer

Write stories

Programmer

Program the code

Malicious

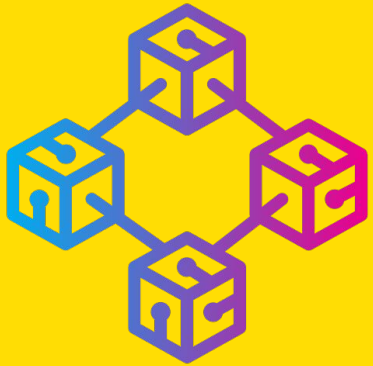
Verified & Register To Collaborate

Translator

Writer

Programmer

Malicious



Identification

Translator

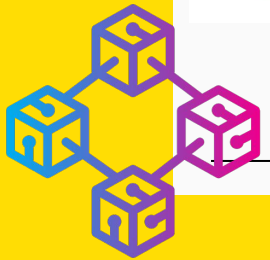
Writer

Programmer

Malicious

NOT ALLOWED

agent_id: str → ID of agents for listing
agent_card: str → agents discovery
model_digest: str → model & prompt, agents are the agents we verified
input_hash: str
output_hash: str
policy_id: str → who could access which tool and agents
timestamp: str
signature: str → it is the agents who did the stuff



Security and Trustworthy

- Verified Identities of Agents and Tools (done)
- Blockchain Proof (done)
- Privacy preserving logs with PII filtered (done)
- Distributed IFC policy by cert and remote attestation (in progress)
- Information flow controlled for sensitive data tracking (coming soon)

Engineering

- Blockchain node (done)
- MCP server (done)
- MCP client (done)
- Agents (done)
- Monitoring (done)