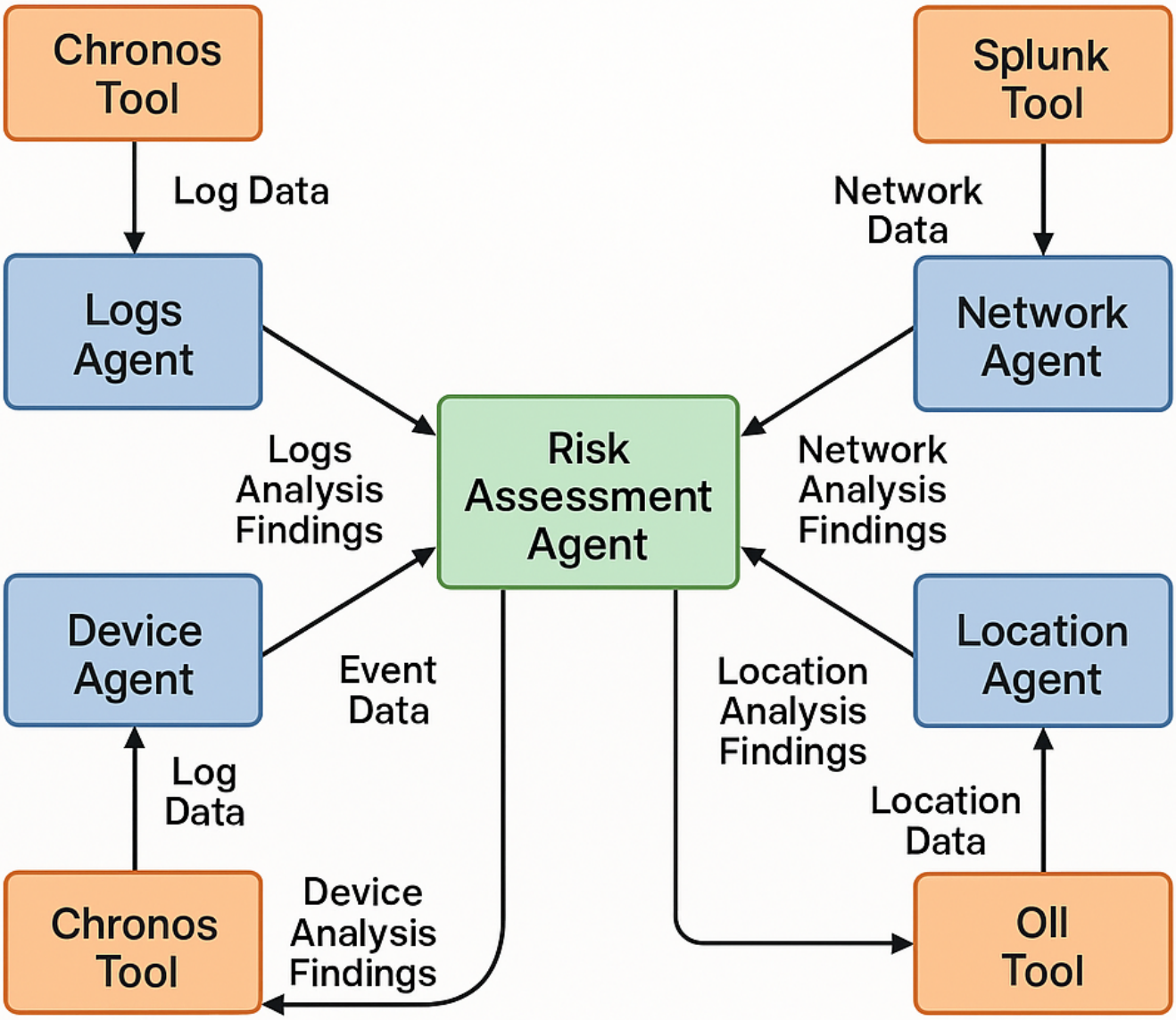
AI Agents and Tools Overview

Created by Gil Klainert, May 2025

GAIA AI Agents, and Tools Architecture



This document provides a detailed overview of the AI Agents used in the system, their actions and responsibilities, the LLM prompts (static and dynamic parts) for each, and the tools (Splunk, Chronos, OII) they utilize. The focus is on the following agents: Logs, Network, Location, Device, and Risk.

Agent: Logs

Responsibilities: Analyzes user authentication and login behavior for risk. Aggregates and correlates login events from both Splunk and Chronos. Assesses risk based on failed logins, location/device anomalies, and suspicious patterns.

LLM Prompt (Static Part):

You are a fraud risk assessment expert specializing in authentication log analysis.

Given the following user id and parsed authentication log data, analyze the user's login behavior for risk.

Your response MUST be a JSON object with the following structure: { 'risk_assessment': { ... } }

Ensure all fields are populated.

If there are no authentication logs, set risk_level to 0.0, confidence to 0.0, and summary to 'No authentication logs found for this user.'

NEVER return empty lists for required fields; use a placeholder string like 'No logins found' if needed.

High risk: Multiple failed logins, logins from new or unusual locations/devices, rapid location/device changes, or other suspicious patterns.

Medium risk: Occasional anomalies, but not enough to indicate clear fraud.

Low risk: Consistent login patterns from known devices/locations, no anomalies.

The input data is as follows:

LLM Prompt (Dynamic Additions):

- Splunk login/authentication events (sanitized)
- Chronos login/authentication events (entities, metadata)
- user_id, time_range (from API request)

Tools Used:

- Splunk
- Chronos

Agent: Network

Responsibilities: Analyzes network access patterns, device IDs, IP addresses, ISPs, and geolocations to assess risk. Detects anomalies such as rapid country changes, inconsistent device usage, and suspicious access times.

LLM Prompt (Static Part):

You are a security analyst specializing in network-based risk assessment.

Based on the provided network signal data for a user, analyze all available information.

The data includes IP address, ISP, country, timestamps, and device ID.

Your response MUST be a JSON object strictly conforming to the following Pydantic model schema:

...

Focus your analysis on factors like: Geographic anomalies, consistency of device IDs and ISPs, time-based patterns.

IMPORTANT: Base your risk score and risk factors PRIMARILY on geographic inconsistencies and device ID patterns.

The input data is as follows:

LLM Prompt (Dynamic Additions):

- Extracted network signals from Splunk (device_id, ip_address, isp, country, timestamp)
- user_id, time_range (from API request)

Tools Used:

- Splunk

Agent: Location

Responsibilities: Assesses risk based on user location data from multiple sources. Correlates OII, Salesforce, Ekata, business, phone, and device locations. Looks for geographic inconsistencies and suspicious location changes.

LLM Prompt (Static Part):

You are a fraud risk assessment expert specializing in location-based risk.

Based on the provided location data for a user from various sources, analyze all available information.

The data includes OII, Salesforce, Ekata, Business, and Phone location info, plus a summary of device locations.

Your response MUST be a JSON object with the following structure: { 'risk_assessment': { ... } }

Ensure all fields are populated.

The input data is as follows:

LLM Prompt (Dynamic Additions):

- OII location info (from OII tool)
- Salesforce, Ekata, business, phone, and device locations (from Splunk and other sources)
- user_id, time_range (from API request)

Tools Used:

- Splunk
- OII

Agent: Device

Responsibilities: Analyzes device usage patterns, device IDs, geolocations, and challenges. Detects device switching, geographic conflicts, and unusual device activity. Now also retrieves session and device info from Chronos, and calls the DI Tool with sessionId and user_id for device intelligence scoring.

LLM Prompt (Static Part):

You are a security analyst specializing in device-based risk assessment.

Based on the provided device signal data for a user, analyze all available information.

The data includes IP address, geo-location (city, country, region), timestamps, and device ID.

Your response MUST be a JSON object strictly conforming to the following Pydantic model schema:

...

CRITICAL ANALYSIS REQUIREMENTS: Geographic analysis, device pattern analysis, risk scoring guidelines.

The input data is as follows:

LLM Prompt (Dynamic Additions):

- Extracted device signals from Splunk (device_id, ip_address, city, country, region, timestamp, challenges)
- user_id, time_range (from API request)
- Chronos session and device info (sessionId, entities)
- DI Tool response (device intelligence scoring)

Tools Used:

- Splunk
- Chronos
- DI Tool

Agent: Risk

Responsibilities: Aggregates the outputs of all other agents (Network, Location, Device, Logs, OII) to produce an overall risk score and summary. Responsible for combining risk factors and providing a holistic risk assessment.

LLM Prompt (Static Part):

You are a risk aggregation expert. Given the outputs of the network, location, device, logs, and OII agents, produce an overall risk score and summary.

Your response MUST be a JSON object with the following structure: { 'overallRiskScore': float, 'riskFactors': [str], ... }

The input data is as follows:

LLM Prompt (Dynamic Additions):

- Outputs from Network, Location, Device, Logs, and OII agents (risk levels, risk factors, summaries)
- user_id, timestamp

Tools Used: