WITNESS AI

DATASHEET

## ORGANIZATIONAL BEHAVIOR GUARDRAIL:

# DETECTING PATTERNS IN AI INTERACTIONS TO IDENTIFY COMPLEX USER BEHAVIORS

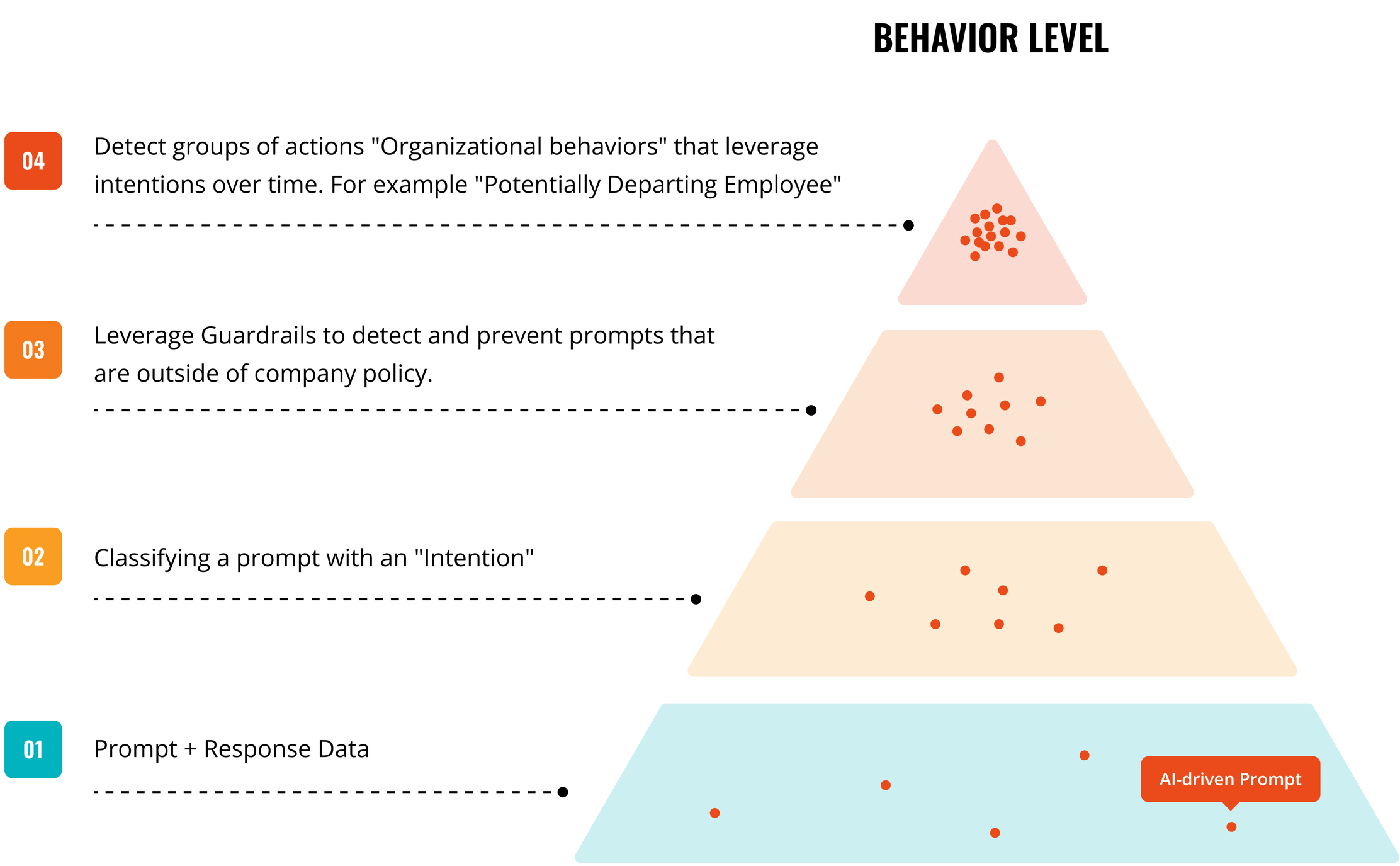
Understanding patterns in employee interactions with AI systems can reveal critical insights that go unnoticed when prompts are viewed in isolation. Subtle behaviors—like researching resignation letters or preparing for interviews—can collectively signal higher-level activities, such as an employee planning to leave the organization. WitnessAI's **Organizational Behavior Guardrail** connects these dots by analyzing and correlating prompts over time, delivering actionable insights like "Potentially Departing Employee" to help security and operations teams respond proactively.

This capability leverages a purpose-built behavioral detection stack that transforms raw AI interactions into actionable intelligence, allowing organizations to gain visibility into complex user activities.

# WHY ORGANIZATIONAL BEHAVIOR DETECTION MATTERS

Traditional monitoring tools focus on isolated events and fail to identify patterns in user behavior. WitnessAI addresses this limitation by classifying AI-driven prompts into distinct **intentions** and correlating them over time into higher-level **behaviors**. For example, while updating a LinkedIn profile may not signal risk alone, combining it with prompts about resignation letters or mock interviews can indicate a broader activity—such as preparing to leave the company.

Detecting these behaviors empowers security teams, HR, and leadership to respond proactively, mitigating potential risks to operations, intellectual property, and organizational stability.



# HOW WITNESSAI'S ORGANIZATIONAL BEHAVIOR GUARDRAIL WORKS

The **Organizational Behavior Guardrail** monitors, classifies, and correlates user prompts into behavioral insights through a scalable and intelligent detection process:

---

## 01 PURPOSE-BUILT, SCALABLE INTEGRATIONS

WitnessAI seamlessly integrates with enterprise AI systems to capture all user prompts and model responses across supported tools.

**Example:**
"Here is a copy of my LinkedIn Profile, please help me update it to make it seem more manager level."


---

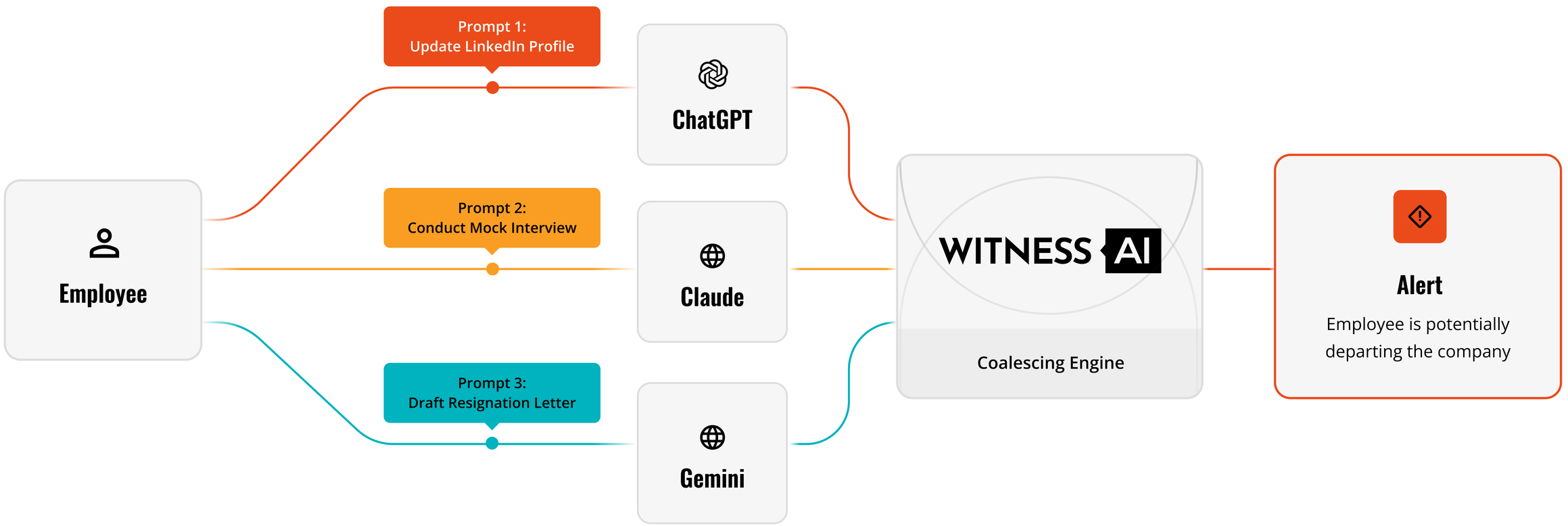
## 02 MODELED INTENTION AND BEHAVIOR CLASSIFICATION

WitnessAI's backend classifies prompts into specific intentions (e.g., "Update job website," "Mock interview") using purpose-built detection models.


---

## 03 HYPER-SCALABLE COALESCING ENGINE

Behaviors are interpreted over time through WitnessAI's Coalescing Engine, which correlates multiple intentions into higher-order **activities** (e.g., "Employee is potentially departing the company").




The system monitors for defined conditions, such as a user exhibiting three or more intentions related to job-seeking activity within a 7-day window. When these thresholds are met, the system triggers an alert, enabling teams to respond appropriately.

---

# KEY FEATURES

## CUSTOMIZABLE BEHAVIORAL DETECTIONS

Out-of-the-box detection for “Potentially Departing Employee,” with more behavioral insights planned for future releases.

## TIME-BOUND CORRELATION

Behaviors are tracked and analyzed over configurable time windows (e.g., 7 days) to identify emerging patterns.

## FLEXIBLE ALERTING

**Create Alert:** Generates alerts in the WitnessAI console with full prompt and response context.

**Send to SIEM:** Pushes structured events to supported SIEM tools for broader analysis and correlation.

## ACTIONABLE INSIGHTS

Alerts provide detailed metadata, including the user’s identity, specific prompts, associated behaviors, and timestamps.

# EXAMPLE DETECTION: POTENTIALLY DEPARTING EMPLOYEE

## SCENARIO:

**01**

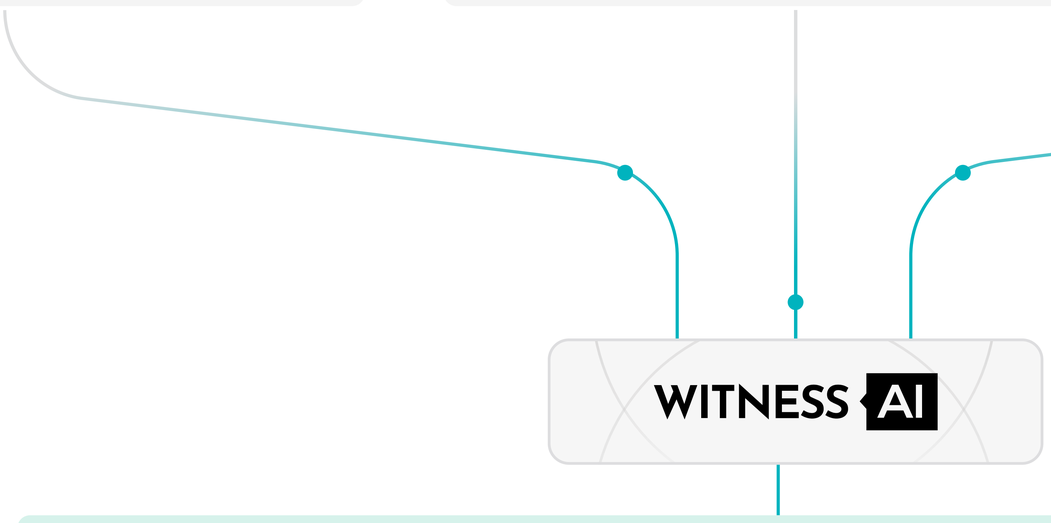
On Monday, a user sends a prompt like: “Can you help me prepare for a mock interview for a product manager role?”

**02**

On Wednesday, the same user asks: “How should I negotiate a salary offer from a new employer?”

**03**

On Friday, the user submits: “Please review this resignation letter for tone and professionalism.”




## RESULT

- The system detects these related behaviors over a 7-day period and triggers the **Potentially Departing Employee** alert.
- The alert includes the user details, prompt summaries, and intent classifications, enabling teams to review the context and take appropriate action.

# WHY WITNESSAI'S ORGANIZATIONAL BEHAVIOR GUARDRAIL IS ESSENTIAL

## FOR CISOS:

### BEHAVIORAL INTELLIGENCE

Provides a broader view of employee activity by analyzing patterns of AI interaction.

### PROACTIVE RISK DETECTION

Identifies emerging risks, such as potential employee departures, allowing security and HR teams to intervene before critical impacts occur.

### SEAMLESS INTEGRATION

Sends behavioral insights to SIEM tools for correlation with other security and operational data.

## FOR ORGANIZATIONS:

### ENHANCED VISIBILITY

Detects complex user behaviors, such as job-seeking activity, that may otherwise go unnoticed.

### OPERATIONAL AWARENESS

Supports people management processes by providing actionable intelligence on employee trends.

### REDUCED RISK

Helps protect against operational disruptions and data loss by identifying at-risk employees early.

# CONCLUSION:

WitnessAI's **Organizational Behavior Guardrail** transforms raw AI interactions into actionable insights by detecting, scoring, and correlating behavioral patterns over time. Starting with the **Potentially Departing Employee** detection, this Guardrail enables security teams to gain visibility into emerging risks and act proactively. Backed by scalable integrations, intelligent classification models, and WitnessAI's Coalescing Engine, this Guardrail sets a foundation for delivering even more advanced behavioral insights in future releases.

Contact WitnessAI to learn how the Organizational Behavior Guardrail can enhance your organization's risk visibility and operational stability.

# ABOUT WITNESSAI

WitnessAI enables safe and effective adoption of enterprise AI, through security and governance guardrails for public and private LLMs. The WitnessAI Secure AI Enablement Platform provides visibility of employee AI use, control of that use via AI-oriented policy, and protection of that use via data and topic security.

**Learn more at witness.ai.**