Remediating Systemic Privacy Incidents

Sam Havron
Privacy Engineer

David Huang
Privacy Engineer

Meta

Outline


What are systemic privacy incidents?


Identifying incident clusters


Analysis


Remediation

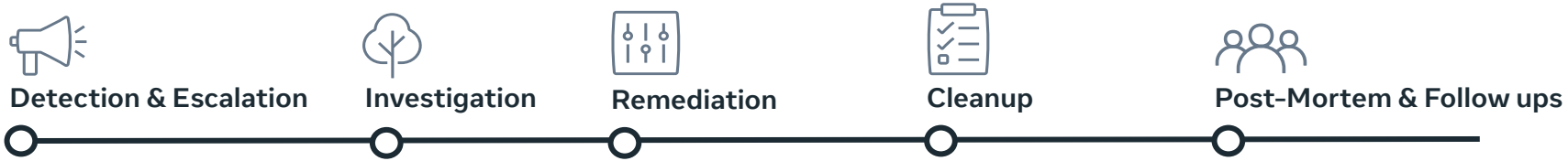

Regressions

What is a privacy incident?

Something is broken that is negatively affecting user privacy

Privacy vulnerabilities largely treated the same as incidents

Lifecycle of a privacy incident



Systemic privacy incidents

*A **group** of privacy incidents with similar root causes or outcomes on users*

Example: a leading smart home company has identified a series of incidents where its robot vacuums are not respecting user consent for processing sensor data

Systemic privacy incidents

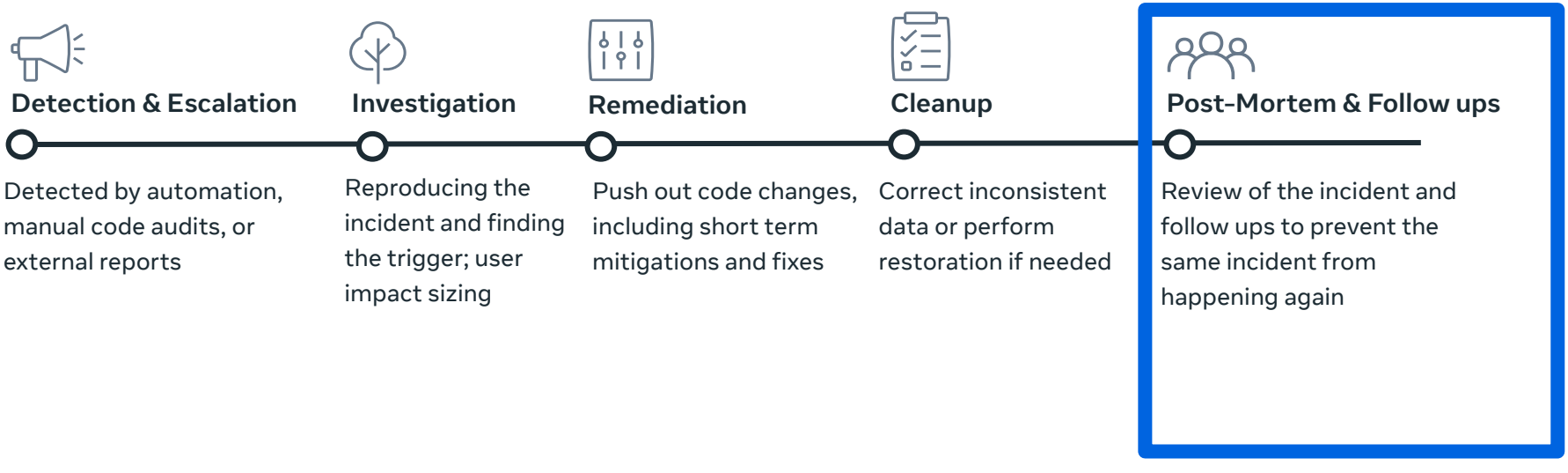
*A **group** of privacy incidents with similar root causes or outcomes on users*

Example: a leading smart home company has identified a series of incidents where its robot vacuums are not respecting user consent for processing sensor data

Root Cause Patterns:

- Consent control implementation is fragmented across robot models
- Pre-consent and opt-out users not segmented in data pipelines
- User controls on the mobile app become out of sync with the server (e.g., stale cache, optimistic updates)

Lifecycle of a privacy incident



Why don't Post-Mortems prevent systemic incidents?

Why don’t individual Post-Mortems prevent systemic incidents?

Post-Mortem & Followups

Review of the incident and follow ups to prevent the same incident from happening again

Missing systemic patterns across incidents due to low visibility of similar issues

Ad-hoc prevention lacks standardization

Follow ups aimed at fixing deeper root causes can be deprioritized due to lack of perceived impact

Why don’t individual Post-Mortems prevent systemic incidents?

Post-Mortem & Followups

Review of the incident and follow ups to prevent the same incident from happening again

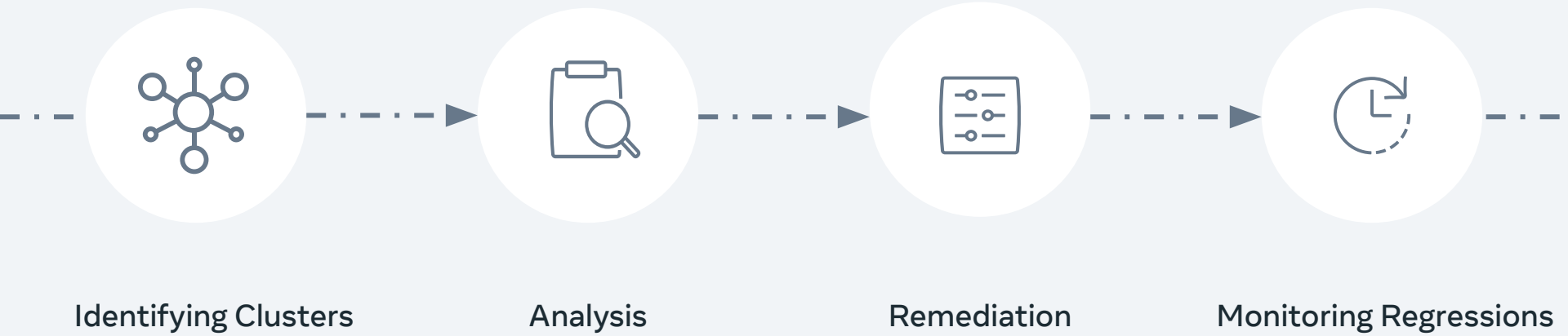
Missing systemic patterns across incidents due to low visibility of similar issues

Ad-hoc prevention lacks standardization

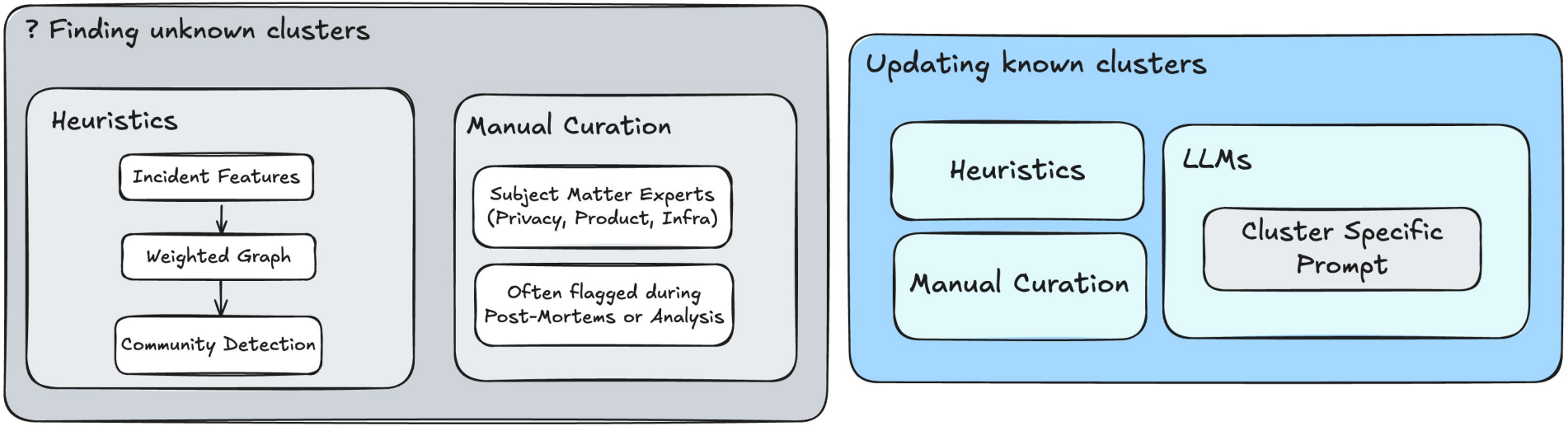
Follow ups aimed at fixing deeper root causes can be deprioritized due to lack of perceived impact

Preventing systemic incidents requires a **comprehensive approach** to address gaps in individual incident handling

Remediating Systemic Privacy Incidents



Identifying Privacy Incident Clusters



Methods

- No one-size-fits-all approach
- Prefer minimizing false negatives to ensure we don't miss potential regressions

Identification: LLMs

Zero-Shot Prompt

Incident context (RAG)

<Incident Summary>
<Trigger details including code references>
<Mitigation details including code references>
<Incident Report, Comments>

➕

System prompt (structured output)

==> Instructions: Given the above incident context, please provide an answer to the query below. Your answer should include:

* Result: 'True' or 'False' if the incident matches the pattern described in the query
* Confidence Level: ‘High’, ‘Medium’, or ‘Low’
* Reason: A detailed explanation of your result

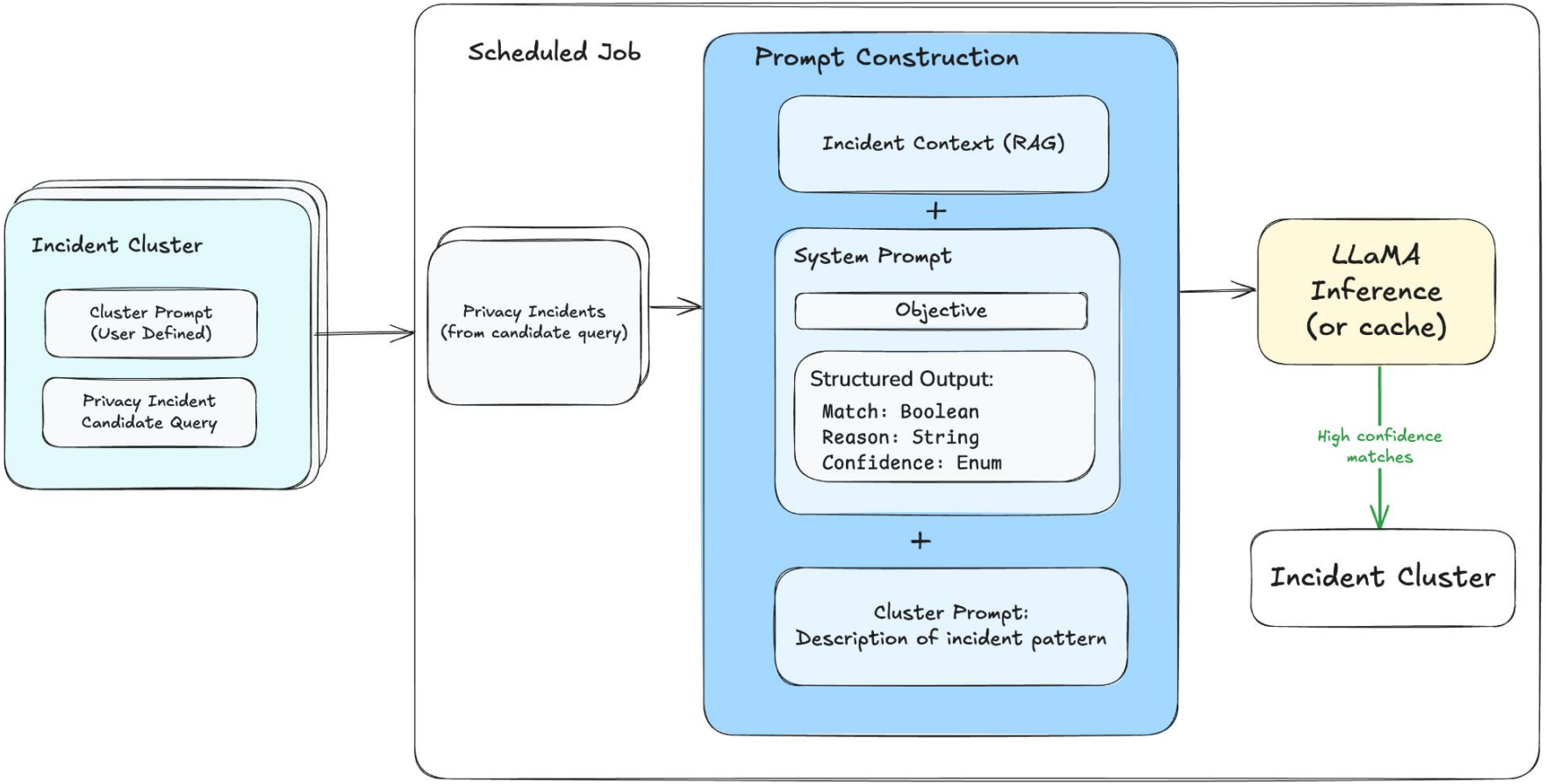
The query is:

➕

Cluster prompt (user defined)

User consent for robots to collect and process sensor data are not respected. The incident trigger is typically related to caching issues, incorrect server handling, or missed consumption by a product surface.

Identification: LLMs



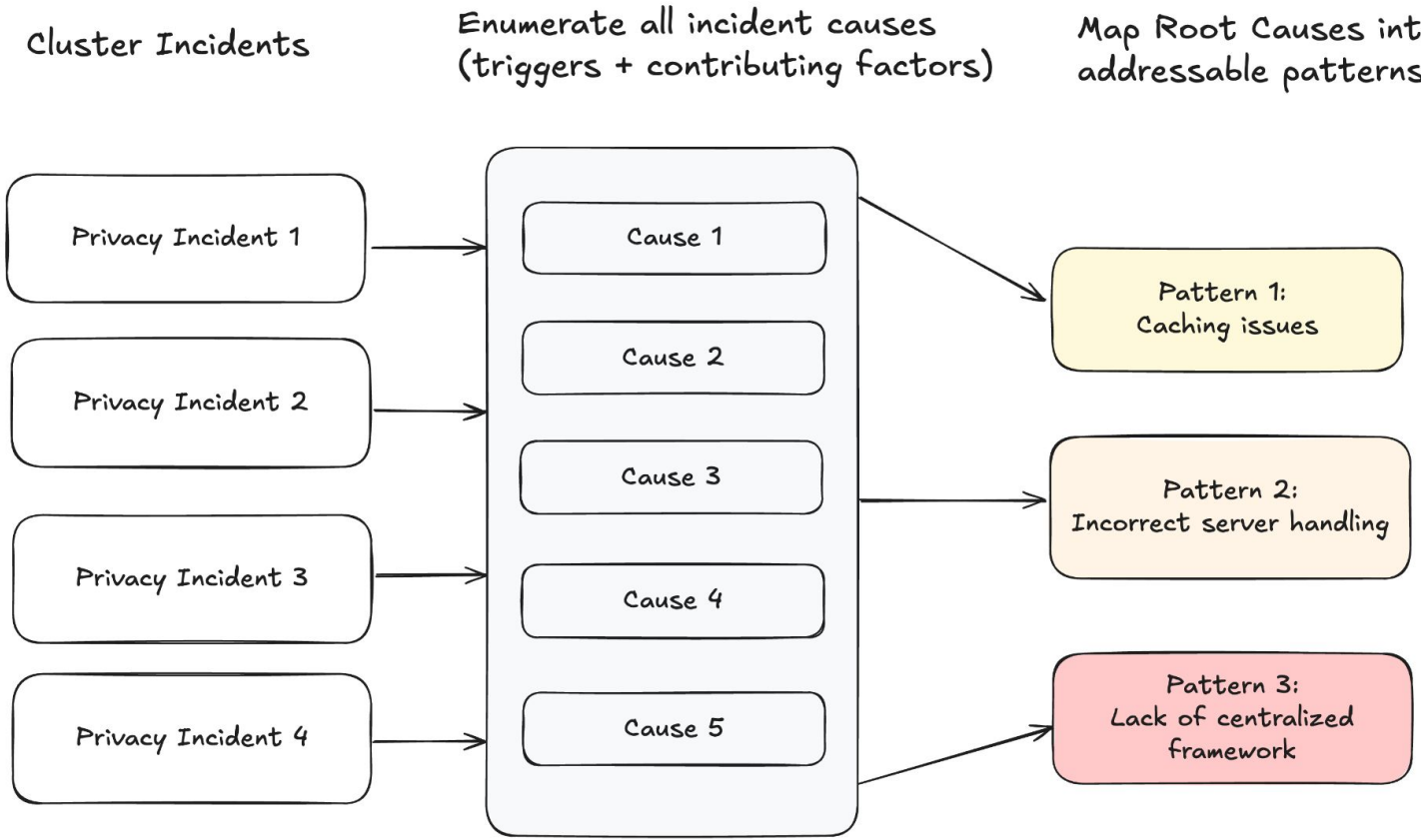
Prioritizing Clusters for Analysis & Remediation

We don’t have infinite resources

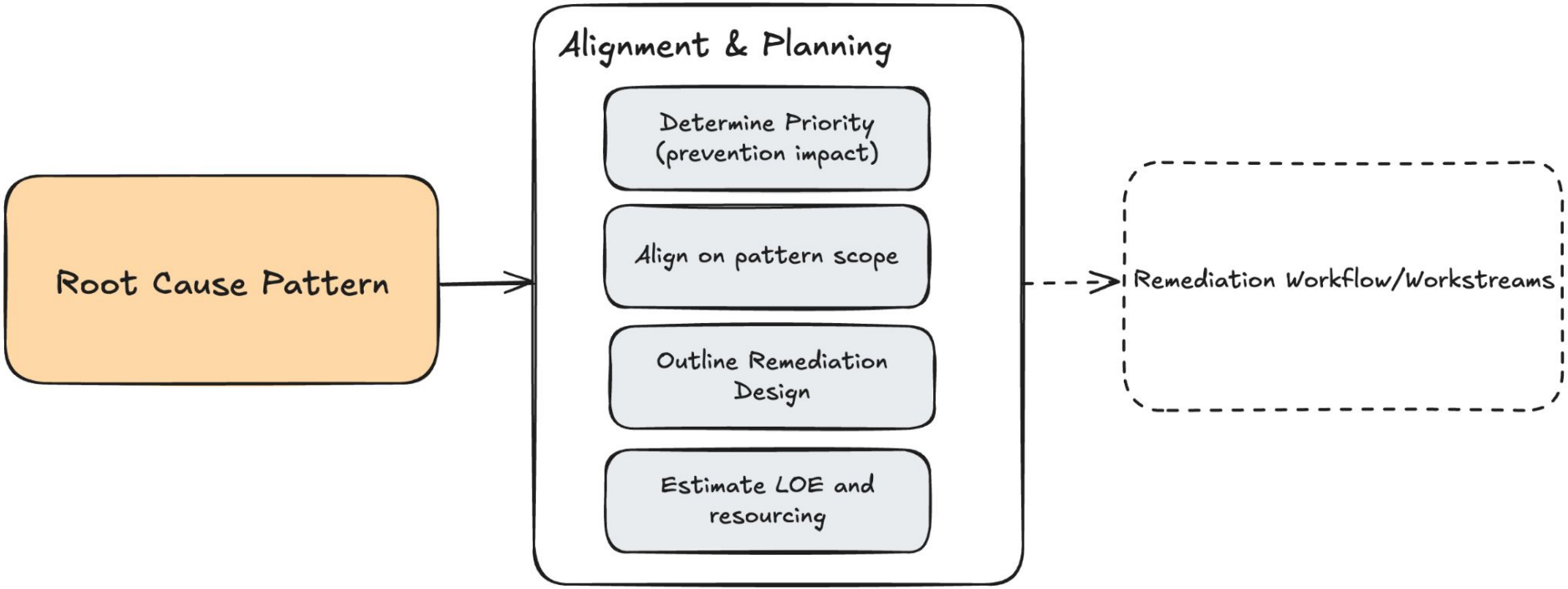
Cluster priority scores based on sum of member incident factors including data sensitivity, user impact, and regulatory requirements

Partnering early with product and infrastructure teams

Analyzing Privacy Incident Clusters

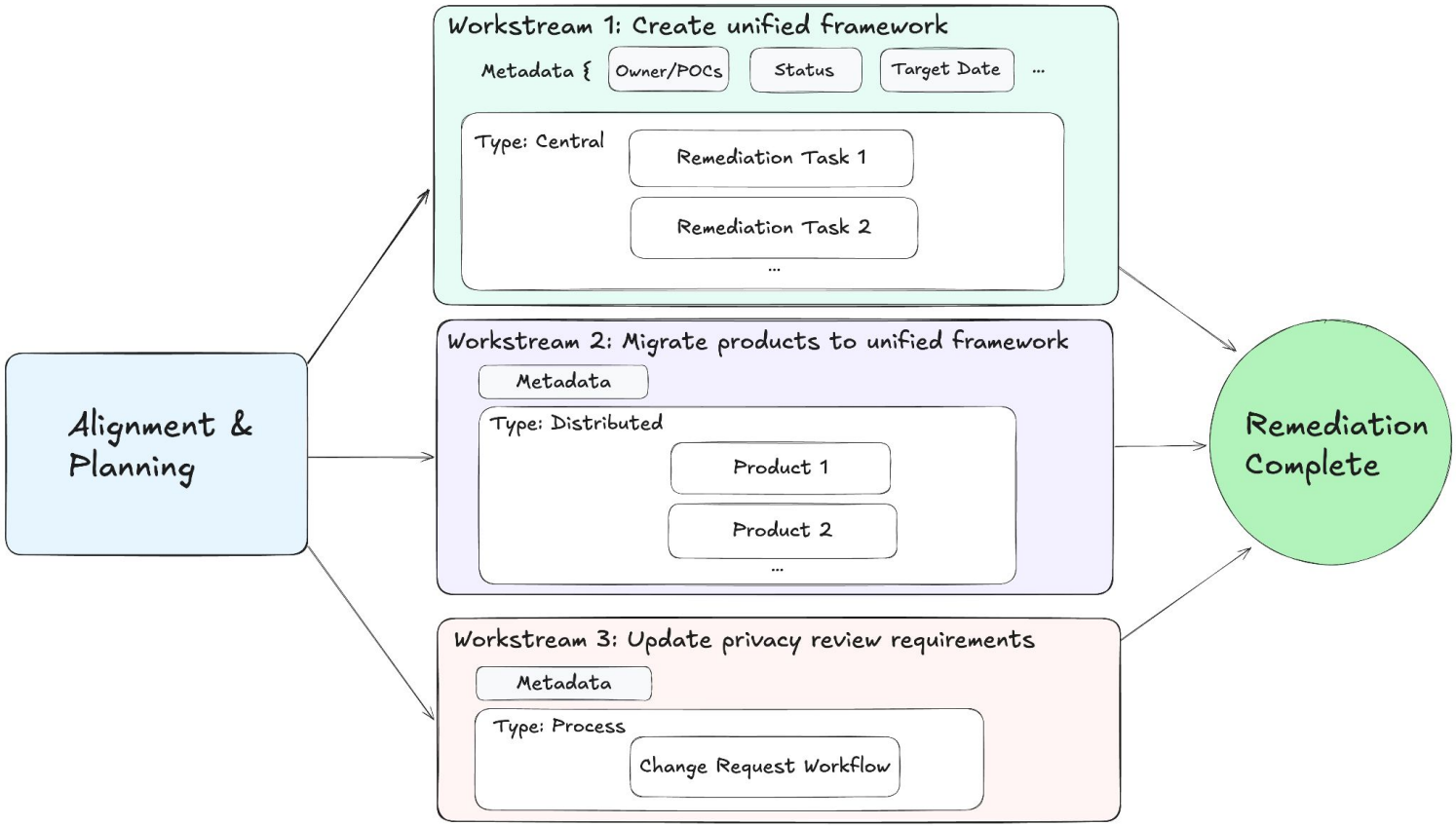


Alignment & Planning



Remediation

Implementation & Monitoring



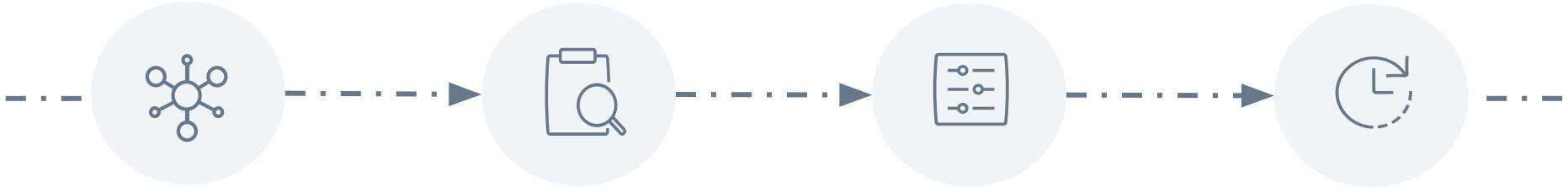
Regression Monitoring

New privacy incidents continuously evaluated with heuristics and LLMs against existing clusters. Matches on completed clusters are potential regressions

Alert cluster owners to confirm regression and review Analysis & Remediation for gaps

Regression monitoring as a lagging metric for effectiveness of remediations

Remediating Systemic Privacy Incidents



Takeaways

- Privacy incidents can share similar root causes or outcomes
- Continuously identifying clusters is critical for designing durable remediations
- Not all remediations are equal

Advice on creating a similar system from scratch

- Start with the basic process and build tooling over time
- Ad hoc and lightweight approaches to identification and analysis

  For instance, group similar incidents in the same post-mortem review
- At smaller scales, some parts of the systemic incident lifecycle may be easier to navigate (smaller codebase, less teams to align). Some parts could be more difficult (prioritization)

Q&A

Meta