**Claude Breakdown: Chunk 10 — Learning Governance, Category Merging, and SOP Evolution**

**✅ Scope**

Define how ClubOS V2 evolves over time: how Claude proposes new logic, how categories are split or merged, and how AI’s learning is human-guided, audit-logged, and scoped to only high-quality signals. Prevents logic sprawl and maintains trust.

**🔁 Cross-Referenced Corrections**

| **Early Plan** | **Final Decision** |
| --- | --- |
| All messages used for AI learning | ✅ Only admin-approved or tagged threads contribute to LLM learning |
| Static SOP categories | ✅ SOP categories can evolve, merge, or branch based on message patterns |
| No pattern summarization | ✅ Claude can propose new categories based on ambiguous thread clustering |

**🔧 Dependencies**

* Claude for similarity clustering + SOP suggestion
* Weekly background job for log + SOP review
* Feedback dashboard interface
* SOP version store with merge/edit capabilities

**📋 Tasks**

**[10.1] Learning Permissions and Filtering**

AI learns only from:

* Threads marked flag\_for\_learning
* Operators with learning-enabled roles
* Interactions where SOP was followed and outcome confirmed

All others are archived for reference but not used as training signal.

Operator tagging (UI):

“✅ Use this phrasing in future”

“📌 Great outcome, promote this logic”

Backend stores:

{

"type": "learning\_tag",

"linked\_thread": "msg-9012",

"linked\_sop": "booking-011",

"by": "admin\_mike",

"confidence\_boost": 0.2

}

**[10.2] Category Merge + Creation**

When Claude detects:

* High overlap across SOPs (e.g. cancellation policies, promo edge cases)
* Or a growing pool of ambiguous messages

It proposes:

* Merge SOPs into one canonical rule
* Create new subcategory (e.g. Calibration, Simulator Prep, Events)
* Re-tag old SOPs to new structure

Admin review modal:

“These 19 messages appear to describe warm-up/cooldown processes.

Would you like to create a new SOP category?”

If approved:

* Claude reindexes SOP vector space
* LLM prompt condition is updated to include new route

**[10.3] SOP Change Summarization + Digest**

Claude auto-generates:

* Weekly summary of merged rules
* Suggested deprecations
* Newly created SOPs from logs

Each change is logged in change\_log:

{

"change\_type": "category\_merge",

"merged\_sops": ["booking-012", "booking-018"],

"summary": "Unified last-minute refund logic",

"status": "approved",

"approved\_by": "admin\_mike"

}

Digest is available in dashboard or Slack/email summary.

**[10.4] Preventing Logic Drift**

To avoid the AI learning false patterns:

* Claude flags outlier threads (“This message deviates from known SOPs”)
* High-volume SOPs are auto-reviewed if override rate is high
* Operator can lock SOPs:

{

"mergeable": false,

"allow\_override\_learning": false

}

**[10.5] Claude-Only Learning Separation**

Claude’s knowledge evolution is scoped to:

* SOP logic
* Categorization structures
* Pattern recognition

GPT-based LLM handles:

* Live routing
* Actions
* Replies

Separation ensures Claude doesn’t learn customer phrasing or human behavior directly.

**✅ Expected Behavior**

* ClubOS evolves only from trusted, intentional inputs
* Claude handles all structural improvement proposals
* Operators control whether logic grows or consolidates
* System scales in clarity, not complexity

Proceed to Ticket system Chunk 11