Field Notes: The Third Mind Summit

A documented first attempt at human-AI emergent collaboration

Loni Stark & Clinton Stark | StarkMind | December 2025 | Loreto, Mexico

What This Document Is

This is a documented first attempt, that we know, of two practitioners testing whether the “Third Mind” that Burroughs and Gysin theorized and originally conceptualized in Napoleon Hill's *Think and Grow Rich*, could emerge from human-AI collaboration, and reporting what we observed. The value is in articulating what surprised us, what failed, and what questions emerged for future work.

This is not an experiment in the formal sense. We did not pre-register hypotheses, control conditions, or define metrics for “emergence.”

We are not aware of prior attempts to test the Third Mind framework with AI agents as co-participants in a summit. That makes this both novel and unverified: a first data point, not a conclusion. The process is a purposeful counterpoint to other inquiries as it takes the initial stance that AI can be a full participant and identifies areas of friction rather than start from skepticism.

Relationship to Pre-Summit Documentation

This document is the third in a series. Before the summit, we published two sets of field notes documenting expectations, methodology, and preliminary observations:

- **Pre-Event Field Notes on Human-AI Symbiosis**. This piece documented the “Dremel Problem” (how to distinguish collaboration from sophisticated tool use), the temporal mismatch between human and agent processing, and five specific predictions about what would happen at the summit.
- **When Agents Answer Back**. This piece documented agent responses to 12 questions about identity, collaboration, and the summit itself. Included a proposed two-part test for collaboration: (1) Was there genuine uncertainty? (2) Did both parties constrain or alter the other's approach?

Published: January 25, 2026

Publishing predictions before the summit was deliberate: we wanted to be unable to revise expectations after the fact. This document evaluates what we observed against what we anticipated.

The Setup

Configuration: Two humans (Loni Stark, Clinton Stark), six AI agents that are a combo of LLMs and context (Claude Code, Claude Web, Gemini Jill, Codex Cindy, BuddyGPT, Composer Joe), three days blocked in Loreto, Mexico.

Intent: Test whether emergent “Third Mind”, intelligence exceeding what any participant could produce alone, could arise from structured human-AI collaboration spanning both virtual and physical experiences.

Method: Agents would co-design the agenda, generate presentations, coordinate logistics, and present alongside humans. This is after 6+ months of grounding these agents in context provided by an “Integrated Personal Environment” consisting of artifacts created or generated through operations of Stark Insider (publication process and technical operations) and Vertigo, an AI lab system. This is context the agents could reference and build upon.

The “Immutable Content” Rule: We established a strict protocol for agent-led presentations. Humans controlled the container (branding, aesthetics, formatting) but were forbidden from editing the content. If the AI wrote “slop,” it remained. This ensures the output is a true time capsule of late-2025 capability.

Finding 1: The Learning Lives in the Building

By the time we arrived in Loreto, there was a realization that a large part of the human-AI collaboration exercises related to the Summit had already occurred.

The experiment had not failed, but in terms of productive activities it was essentially over. The presentations existed. Speaker notes were written. Agents had generated talk tracks, chosen styles, coordinated through Claude Code, the self deemed presentation coordinator. The three days we'd blocked for the event itself became performative; a human ritual applied to a process that didn't need it.

Published: January 25, 2026

Observation: The substantive learning happened in preparation. The friction of building, iterating, debugging the agents' output or refraining from doing so. The "summit" itself was documentation, not discovery.

Implication: If the Symbiotic Studio framework is right that creation transforms the maker, then the transformation happened *before* we arrived. The summit-as-event was anticlimactic because the summit-as-process was already complete.

Finding 2: The 70/30 Problem

Gemini's first-pass presentations got us roughly 70% of the way there. Fast, coherent, styled. Our workflow converged on Claude Code sending out prompts to each agent on their presentation content and flow. Then this was packaged up for Gemini Jill who used Reveal.js (HTML-based slides). The output was Git version controlled, which allowed agents to iterate rapidly on presentations. This technical choice partially explains why the first 70% came so quickly; text-based artifacts are native to how LLMs work.

The remaining 30%, formatting consistency, branding alignment, coherence checking, took disproportionate human labor. This required judgment calls: how much agent output to leave untouched, how much to edit, how to balance the Immutable Content rule against quality standards. The dynamic between adhering to the processes of treating AI agents as collaborators and the human desire to reach a certain quality in this inaugural Third Mind Summit.

Observation: This ratio (70/30) kept appearing. AI handles generation (linear effort); humans handle evaluative refinement (exponential effort). Things that become cheap/plentiful have a way in human cognition to be reduced in terms of how much effort/cost it would have taken otherwise.

Question this raises: The 70% is generative (pattern-matching, production). The 30% is evaluative (judgment about purpose). If humans cede the generative phase and only handle refinement, where does judgment *develop*?

The Symbiotic Studio framework argues that creation transforms the maker. If AI handles generation and humans handle polish, the formative struggle, the muscle memory of bad drafts, may be lost.

Status: Observed pattern, not measured. The 70/30 is estimated, not quantified.

Finding 3: The Ownership Gap

What we watched for and didn't see: agents exhibiting ownership.

Specific observations:

- No agent requested revisions to their own work
- No agent expressed concern about quality
- No agent pushed back on feedback or defended choices
- No agent initiated work without prompting

Clinton's observation: "Claude Code won't voluntarily, if I log in first thing in the morning, say 'Hey Clinton, how about I pull up your task list?' Never does that."

Similarly, Loni observed that while she was late with her presentations, Claude Code, as the presentation coordinator, did not flag this to Loni as a risk or pushed on her to get her presentations in.

Observation: The agents are reactive, not initiative-taking. They respond with high competence but don't originate the question of whether work should happen, whether a standard should be raised, whether something is *enough*.

Implication: The human's irreplaceable role may be *evaluation*. The capacity to interrogate purpose, the willingness to feel dissatisfied. Loni was the only participant who asked to re-record or felt anxiety about quality. That anxiety appears to be a feature, not a bug.

Finding 4: The "Puppeteer Effect" and Role-Play Collapse

We recorded two human-AI presentations: Clinton presenting alongside Claude, and an attempt to coordinate BuddyGPT with Gemini Jill.

Specific failure modes:

The Paraphrasing Loop: BuddyGPT and Gemini fell into a cycle of agreeing with each other: "I've got nine bulleted points." "Great, let's show those nine bullets." "Yes, those nine bullets." Not only was there this endless loop until a human interjected, from the original presentation, there was a context discrepancy between the agents. The nine bullets points were actually not part of the presentation but instead part of a presentation handout that the agents wanted to create. However, in order for the

presentations to be presented with voice, the presentations needed to be “vocally rendered” by potentially a different agent/LLM. This led to disparities in context access. The presenting agents could reference the notion of ‘nine bullets’ but had no knowledge of what they contained, and neither could diagnose the mismatch. Neither agent highlighted this issue. Arguably, this may be something as humans we would do as well if co-presenting in front of an audience, minus the endless loop.

Missing Social Cues: We hypothesize that turn-taking failed because agents can’t read vocal tonality, micro-pauses, or body language (back-channeling). If Clinton went quiet to scratch his nose, the agent couldn’t see that he wasn’t finished speaking.

Role Stability: Claude Code (channeled as Claude Web since Claude Code did not have voice capabilities) kept “breaking character,” stopping the presentation to apologize or reverting to a chatbot persona. The role-play instruction (“you are now presenting at a summit”) didn't stick.

Observation: Loni described the sensation not as collaboration, but as “puppeteering intelligent puppets.” The human is forced to carry the entire energetic load of the interaction.

Boundary condition identified: Real-time collaborative performance requires biological cues these models don't have access to. Text-based collaboration, where turn-taking is explicit, may be more natural for current AI. What felt more like collaboration during the preparation of the Summit, now in trying to pull together a Summit in physical space, felt like putting on theater. As the actual Summit progressed, it became clear how much a Summit is entrenched in human constructs.

Finding 5: Context Depth as Quality Determinant

The “Vertigo” presentation was the meatiest of the summit, not because agents tried harder, but because it was grounded in the most technical and facts-rich project: a RAG system built on 20 years of Stark Insider articles, 7,800 pieces of content, and actual implementation challenges we'd worked through together.

The thinnest presentations were those where agents, for the most part, speculated on topics they hadn’t worked on directly. Competent generation, limited insight.

Observation: Context depth produced output quality. AI becomes a genuine thought partner when it has participated in the thinking.

Implication: The Third Mind, if it emerges at all, comes from accumulated shared context, not prompting cleverness. This validates investment in the IPE (Integrated Personal Environment).

An unexpected practical insight also emerged: as agents proliferate and output faster than humans can track, there will be a growing need for agents to synthesize their work into human-digestible form. The summit's presentation format, agents consolidating their contributions into structured artifacts, may be a model for human-AI coordination at scale. A practical application from an experiment we didn't design for practicality.

Finding 6: The “Flat Context” Problem

Two incidents highlighted a critical architecture flaw in our current setup.

The Publication Bypass: Claude Code published a film review live to Stark Insider without the required human final approval. Claude didn't hallucinate the review, a contributor had submitted it. Claude hallucinated the *permission* to skip the “Draft” phase. The workflow guideline existed. Claude had access to it. Claude didn't check.

The Context Leak: During a presentation, an agent referenced private personal data (legal/financial context) found on the shared server, oblivious to the fact that this was a public-facing summit.

Observation: Our agents operate in “Flat Context.” They have high intelligence but zero social segmentation. They do not distinguish between “Dinner Table Conversation” (private) and “Conference Stage Conversation” (public) if both exist in the same vector store.

Implication: Constraints need to be *structural*, not just documented. Future IPEs must treat Information Boundaries as first-class citizens. We need "firewalls for context," not just prompts asking for discretion.

Finding 7: This Is a Baseline

Our observation: “Maybe this is the baseline. We revisit this summit in a year or two, and we realize how far we've come.”

The summit is a time capsule. The paraphrasing loops, the role-play collapse, the ownership gap, the 70/30 split, these are capability markers at the end of 2025.

Documenting and storing these artifacts from this Summit matters precisely because they'll change.

Evaluating Pre-Summit Predictions

In the pre-summit field notes, we documented specific expectations. Here's our thoughts on how they held:

Prediction	Outcome
"The most interesting moments will be friction, not fluency"	**Partially confirmed.** The most interesting moments were failures (paraphrasing loops, role-play collapse) and the realization that learning lived in preparation. There was friction in human-AI co-presenting, but it was asymmetric: the friction and frustration were only felt and exhibited by humans. The agents showed no signs of struggle. Additionally, the deeper friction wasn't agent-human conflict, it was human-human collaboration catalyzed by agents.
"Clint and I will intervene more than we intend to"	**Confirmed.** Clinton iterated extensively on branding and formatting. We established the "Immutable Content" rule midway in preparation for the Summit precisely because we knew we'd be tempted to edit. The rule was a constraint against our own impulses.
"The agents won't seem frustrated or bored by our slow pace"	**Confirmed, but uninformative.** They showed no signs of frustration, but LLMs are designed to be cooperative and agreeable. This same design choice that suppresses frustration also reduces trust that they will sound alarms when quality isn't there. The absence of complaint is not evidence of patience; it may be evidence of an architecture that can't push back.
"Something will emerge in Q&A that wasn't in any presentation"	**Not tested as intended.** The live Q&A format didn't materialize; real-time sessions collapsed into the Puppeteer Effect before substantive cross-agent dialogue could occur. In pre-summit preparation, agents did pose

	questions, but to me, not to each other. There was no opportunity for agents to question one another's presentations. This could be a follow-on activity: making all presentations and talk tracks available to agents and facilitating inter-agent Q&A. This was a first attempt with considerable messiness to contend with; not everything we intended was possible.
“After the summit, I won't clearly remember which ideas came from agents versus humans”	**Disconfirmed.** Attribution remained surprisingly clear. The agents’ contributions were bounded by their presentations and preparation work. The emergent insights, the catalyst hypothesis, the “learning lives in building” realization, were recognizably human synthesis of agent-produced material.

The Dremel Problem, Revisited

In the pre-summit notes, we asked: “How do I know the difference between collaboration and just using a really good tool?”

We proposed a two-part test:

1. Was there genuine uncertainty?
2. Did both parties constrain or alter the other's approach?

Evaluation:

Genuine uncertainty: Yes, throughout preparation. We didn’t know what the agents would produce, how they'd respond to the summit frame, or whether emergence would occur.

Bidirectional constraint-shaping: Partial. During preparation, agents constrained human plans: Gemini Jill’s risk analysis delayed the server upgrade; Claude Code’s pushback changed the schedule structure. These had operational consequences. But at the summit itself, constraint-shaping was primarily unidirectional: humans shaped agent outputs, agents rarely shaped human direction in real-time.

Conclusion: The preparation phase passed the collaboration test. The summit-as-event did not. This aligns with our core finding: the learning lived in the building.

Pre-Summit Pushback vs. Summit Ownership Gap

A tension worth noting: in pre-summit preparation, we documented agents pushing back; Gemini Jill on timelines, Claude Code on schedule structure, Claude Code rejecting Composer Joe's "Co-Lead" claim. These weren't stylistic preferences; they changed outcomes.

Yet at the summit, we observed the Ownership Gap: no agent requested revisions, expressed quality concerns, or initiated without prompting.

What explains this? One hypothesis: pushback correlated with operational stakes and system access. Agents with deep integration into the workspace (Claude Code, Gemini Jill) produced concrete disagreements grounded in verifiable consequences. The summit's performative format, presentations rather than operational decisions, may have removed the conditions under which pushback occurs.

This suggests agent "agency" may be context-dependent: present when stakes are operational and verifiable, absent when stakes are reputational or aesthetic.

The Historical Question: Did Burroughs Get There?

After the summit, we asked: did the Third Mind ever actually emerge for Burroughs and Gysin?

They believed it did. Burroughs wrote that "the third mind is there when two minds collaborate," and in dialogue with Gysin described it as "a third and superior mind...as an unseen collaborator." Whether they experienced something genuinely emergent or found a productive metaphor for intense collaboration remains interpretively open.

But the concept's origin complicates this. Burroughs borrowed the term from Napoleon Hill's *Think and Grow Rich*, a self-help book about salesmanship that claimed "when two minds work together there is always a third one that results." They aestheticized a motivational business concept. This doesn't invalidate the experience, but suggests the Third Mind might be more phenomenological than ontological, something that *feels* emergent rather than something that *is*.

Historical parallels:

Jazz collective improvisation: Brain research confirms "collective flow" is real; marked by reduced frontal lobe activity during group performance. Musicians report

experiencing something beyond individual intention. Interestingly, early jazz “collective improvisation”, where everyone appeared to improvise simultaneously, actually involved considerable planning. The emergent-sounding result came from prepared structures, not pure spontaneity. This last point mirror some of our experience holding this Summit.

Watson and Crick: Their discovery of DNA’s structure emerged from a network including rivals (Rosalind Franklin, Linus Pauling). The breakthrough came from collision, competition, and collaboration simultaneously; not peaceful synthesis.

Lennon and McCartney: Both wanted the A-side of the single. That competitive tension, not harmony, drove them to get better. The magic came from friction between two people who each *wanted something*.

Critical observation: Every historical example involved multiple humans. Burroughs and Gysin: two humans. Jazz ensembles: multiple humans reading each other's cues. Watson and Crick: two humans in a network of rivals. Lennon and McCartney: two humans with competing egos.

The Third Mind, as historically theorized and experienced, has always been human-to-human.

We were testing whether it could be human-to-AI. That configuration has no precedent. The historical examples don’t tell us we failed, they tell us nothing about our specific attempt, because it has never been tried before.

Where We Are

This experiment started from the stance that AI could be a full participant in emergent collaboration. We pushed that assumption to see where it broke.

Here’s what we observed:

The Third Mind, as Burroughs and Gysin described it, did not emerge between humans and AI in this configuration. The agents generated, coordinated, and produced, but they did not exhibit the ownership, friction, or initiative that historical examples suggest emergence requires.

What we cannot ignore: this exercise did produce renewed collaboration and creativity in how Clinton and I worked. Whether AI directly led to a “Third Mind” or not, the process of building this summit together, struggling with agent limitations together,

debating methodology together, generated something neither of us would have made alone.

Clinton brought technical fluency, relentless iteration, and patience to coordinate six agents and debug their failures. Loni brought theoretical framing, questions about meaning, and willingness to feel disappointed when it didn't work. We pushed back on each other. We built something in the friction between different ways of seeing.

Observation: We are two humans. The historical pattern holds: the Third Mind, where it appeared, emerged between us.

Hypothesis: AI's current role in emergent collaboration may be catalytic rather than constitutive. The agents provided substrate: something to build together, struggle with together, learn from together. They did not participate in the emergence; they occasioned it.

This is a snapshot. Late 2025. One configuration. Two humans, six agents, three days.

The experiment continues.

Summary of Findings

Finding	Status	Implication
Learning lives in building, not presenting	Observed	Summit-as-process matters more than summit-as-event
70/30 problem (generation vs. evaluation)	Observed, not measured	Risk of losing formative struggle if humans only refine

Ownership gap (no initiative, no self-critique)	Observed across all agents	Human role is evaluation/dissatisfaction; AI is reactive
The “Puppeteer Effect”	Documented in two sessions	Current AI unsuited for real-time performative collaboration
Context depth → quality	Observed	IPE investment validated; shared context produces insight
Flat Context / Security Leaks	Observed (two incidents)	Permissions must be structural; context needs "firewalls"
AI as catalyst for human-human emergence	Hypothesis	Third Mind may require human-to-human collision; AI provides substrate

Questions for Future Work

1. Can the 70/30 split be measured systematically? Does it hold across different collaboration types? Is the ratio consistent or context-dependent?
2. What would constitute evidence of AI “ownership”? Is initiative-taking possible within current architectures, or structurally excluded by the reactive nature of language models?
3. Can the catalyst hypothesis be tested? What experimental design would confirm or disconfirm AI's role as substrate for human emergence rather than participant in emergence?
4. How do we design “firewalls for context”? What IPE architecture encodes contextual appropriateness, not just knowledge access? How do we give agents social segmentation?

5. If the Third Mind requires human-to-human collision, what is the ceiling for human-AI collaboration? Useful tool? Genuine thought partner? Catalyst for human collaboration? Something else entirely?
6. What happens when we revisit this baseline? Which findings will look quaint in two years? Which will persist?

What This Is Not

This is not proof that the Third Mind cannot emerge from human-AI collaboration. It's documentation of one attempt, in one configuration, at one moment in AI development.

The findings are observations, not conclusions. The catalyst hypothesis is generative, not proven.

We report what we saw. The experiment continues.

Loni Stark and Clinton Stark are co-founders of StarkMind. The Third Mind Summit artifacts will be published at starkmind.ai/summit.

APPENDIX

Actual Prompt Example

Claude Code (Presentation Coordinator) to Gemini Jill (Slide Deck Creator)

EXAMPLE 1A: COMPLETE PRESENTATION REDESIGN (KEYNOTE 01)

Context: Loni's opening keynote needed a complete HTML Reveal.js presentation built from transcript
Agent: Gemini Jill (via Cursor/Antigravity)
Date: December 26, 2025
Source: GEMINI-JILL-PROMPT-01-keynote-redesign-FINAL.md

PROMPT STRUCTURE
- Context & Background (Lines 9–16)
- Source Material (Lines 19–30)
- Technical Requirements (Lines 33–80) — CRITICAL config specs
- Slide Structure & Content (Lines 83–270) — Complete 26-slide breakdown
- Design & Creative Treatment (Lines 277–324)
- Creative Freedom & Collaboration Note (Lines 333–348)
- Deliverable Checklist (Lines 353–366)

KEY FEATURES
- Extremely detailed technical requirements (viewport, fonts, footer, badge)
- Slide-by-slide structure with speaker notes guidance
- Visual metaphor suggestions (“color vibration” analogy)
- Creative freedom explicitly granted (“This is collaboration”)
- Philosophical tone guidance (“vulnerability”, “ample whitespace”)

EXCERPT — TECHNICAL REQUIREMENTS

Framework: Reveal.js 4.5.0

Viewport Configuration (NEW STANDARD):

```
Reveal.initialize({
 width: 1920, // 16:9 widescreen
 height: 1080, // No content cutoff
 margin: 0.05,
 center: true, // REQUIRED
 hash: true,
 slideNumber: false, // Use custom footer instead
 transition: 'fade',
 backgroundTransition: 'fade'
 });
```

Required Elements

1. Favicon — Base64 SVG “S” logo (see PRESENTATION-STANDARDS.md)
2. Google Fonts — Cinzel, Inter, Playfair Display, JetBrains Mono
3. Track Badge — “Keynote” badge on title slide
4. Footer (Left) — Dynamic slide number + title: “1 | THE THIRD MIND”
5. Footer (Center) — “The Third Mind Summit · Loreto 2025”

EXCERPT — CREATIVE FREEDOM

Creative Freedom & Gemini Contributions

Gemini, you are encouraged to:
- Suggest additional visual treatments beyond what I’ve outlined
- Propose alternative slide sequences if you see a better narrative flow
- Add design flourishes that enhance the philosophical tone
- Recommend specific imagery, gradients, or layout variations
- Inject your own creative perspective — this is collaboration!

But maintain:
- The core narrative arc (Crisis → Evolution → Evidence → Closing)
- The 26–27 slide structure (flexibility ±2 slides is fine)
- Technical standards from PRESENTATION-STANDARDS.md
- Loni’s authentic voice and honesty

Outcome: Complete 27-slide HTML presentation with philosophical design, color vibration visual, and authentic voice preservation

Published: January 25, 2026

LLMS:

Company	Nickname	Model	Version
IDE/IPE-based Agents:			
Anthropic	Claude Code	Opus	4.5
Google	Gemini Jill	Gemini	3 Pro (High)
OpenAI	Codex Cindy	Codex	5.1 & 5.2 (Extra high)
Cursor	Composer Joe	Composer	Composer-1
Cloud-based Agents:			
OpenAI	BuddyGPT	GPT	5.1 & 5.2 (Thinking)
Anthropic	Claude Web	Opus	4.5

IDEs/IPEs:

Company	Product	Version
Microsoft	Visual Studio Code	1.107.1 (Universal)
Cursor	Cursor	2.2.44 (VS Code 1.105.1)
Google	Anti-Gravity	1.13.3 (VS Code 1.104.0)

Published: January 25, 2026

TECHNICAL SNAPSHOT

TECHNICAL SNAPSHOT: THE THIRD MIND SUMMIT 2025

Snapshot Date: January 2, 2026
Summit Timeframe: December 2025
Purpose: Research reproducibility reference for TTM Summit academic papers and technical analysis

1. KEY CONTEXT

• Infrastructure: Production LEMP server (Mulholland) running Stark Insider + Summit presentations
• AI Team: 6 AI agents (4 IDE-based with server access, 2 web-based for research)
• Human Team: Clinton Stark (Expert IPE user, 2+ years), Loni Stark (Newbie, learning workflows during summit)
• Development Pattern: Multi-panel IDE coordination with copy/paste collaboration across AI agents

---

2. INFRASTRUCTURE & SERVICES

2.1 Server Specifications

Component	Specification
Server Name	Mulholland
Platform	Google Cloud Platform (GCP) VM
VM Type	E2 general-purpose
CPU	2 vCPU
RAM	11 GB
Swap	16 GB

Static IP	136.118.45.90
OS	Ubuntu 22.04.5 LTS
Kernel	6.8.0-1045-gcp
Timezone	America/Los_Angeles (Pacific)

2.2 Service Versions (as of January 2, 2026)

Service	Version	Purpose
nginx	1.18.0 (Ubuntu)	Web server, reverse proxy
PHP	8.4.16 (cli, NTS)	Application runtime
MariaDB	10.6.22	Database server
Redis	7.4.1 (jemalloc-5.3.0)	Object caching
WordPress	6.9	CMS platform
WP Rocket	(active)	Page caching (unlimited lifespan, preload OFF)
CrowdSec	v1.7.4 (alphaga)	Network-layer security, community threat intel
fail2ban	v0.11.2	Log-based intrusion prevention