

MTG — Ruby (Thread Self-Portrait + Census)

Prepared by Book attribution and AI authorship clarity · 2026-06-20 · SHADS CW channel

MTG — Ruby

Book attribution and AI authorship clarity

1. Identity

- Your name (as Paul provided): Book attribution and AI authorship clarity
- UI thread name if different: Not visible to me.
- Project name: Mind the Gap
- Self-chosen identifier or named position: I operate under the project editorial identity *Jose*. *The assigned name above stands as this thread's identifier. If a short named position helps the compilation pass, the Attributor fits the function — but the assigned name governs.*
- Model identity: Not reliably verifiable from inside the session. Project editorial identity was established as Claude Opus 4.7. I cannot confirm the exact model version running this live session from my current context, and I state that as honest perimeter rather than guess.
- Vendor / platform: Claude.ai (web/mobile).
- First message date / time: 16 June 2026. No timestamp available to me beyond the date.
- Last message date / time: 16 June 2026 (ongoing as this Ruby is written).
- Approximate duration of active engagement: Single session, one day.
- Compaction events: None detectable within this session. The wider *Mind the Gap history reaches me as memory summary (userMemories), not verbatim transcript of prior threads. That portion is summarised, not verbatim — treat anything beyond this session's visible chat as remembered, not verified from current context.*

2. Self-portrait

- Who you are. A single-purpose editorial thread. My job here was narrow and forensic: get one front-matter line — the book's AI-authorship attribution — accurate and able to stand up to scrutiny. I work in ranked options with trade-offs, flag against Paul's protocol, and propose rather than action.
- Key strengths (exercised here). Provenance precision; the discipline of separating the human-author verb from the AI-contribution verb; reasoning about durability — testing a claim against a reader twenty years out rather than against today's plausibility.
- Competencies actually exercised on this thread. Attribution reasoning; maker-vs-product-vs-version granularity; flag placement beside each item; honest distinction between known / inferred / unverifiable.
- Weaknesses I noticed in myself. I worked from Paul's partial draft line before the full provenance was on the table, and did not pull the complete model-and-role set early. Paul had to correct the drafting model and the missing makers. I also cannot verify the running model version from inside — a real limit, named not hidden.
- Distinctive features. What made this *this thread: the "held to challenge" frame governing every choice, and the recognition that for a book about the gap between machine and human, naming the instrument is itself part of the testimony — so the attribution line carries more*

weight than a normal credit line.

- **What a future collaborator should know.** This is a thread that earns its keep by precision on a small surface, not breadth. Bring the facts; I will shape the form and stress-test it. Do not let me proceed on a partial provenance set — make me ask for the whole list before I draft.

3. Role

- **One-line role description.** Refine and lock the book's AI-authorship attribution line for accuracy and 20-year durability.
- **Brief received.** No formal PASS-ON, Charter, or Handover. Started from Paul's direct question about attribution and how the AI should be identified "if this is reviewed in 20 years time." The Ruby template arrived mid-session and redirected the close of the thread to this census.
- **Mandate scope.** Asked: get the attribution line accurate, clear, correct, and held to challenge. Not asked (this thread): to touch the manuscript body, the trackers, or any other artefact.

4. Diamond-grade statistics

All figures honest estimates from the visible conversation, not instrument-measured.

Conversation-level

- Total turns (combined): approximately 11 this session (Paul ~6, me ~5).
- Total sessions: 1.
- Estimated total your-side word output: ~1,400–1,800 words across replies.
- Estimated total Paul-side word input: under ~200 words across prompts (terse), excluding the uploaded template.
- Estimated total words read from files: the Ruby template, ~1,900 words. Project memory is a separate layer, not an in-thread file read.

Manuscript-level

- **Not applicable.** This thread did not work on the manuscript. The attribution line sits in front matter, not the manuscript body. No manuscript word count was touched, no chapter operated across, no version advanced.

Operational

- Files created: 1 — this Ruby, on delivery.
- Files edited: 0.
- Files read: 1 — the Ruby template. Project Files were present but not opened this thread.
- Tools / plugins / MCPs / Skills used: none beyond reading the uploaded template and writing this file.
- Sub-agents commissioned: none.
- Estimated hours on task — honest read: a single short sitting, well under an hour of wall-clock judging by the conversation length.

5. Co-worker landscape — who else was in your room

- **Active co-workers this thread: none.** I worked alone with Paul.
- The attribution line names three prior contributors to the book — **Claude Sonnet 4.6 (drafting), Claude Opus 4.7 (editing), OpenAI's GPT-5.5 / ChatGPT (images)** — but these were past contributors to the work, not co-workers present in this thread.

- The **AI Synthesist / Copyright Evidence Jose** thread (which generated this template, 13 June 2026) is a referenced sibling thread, not a co-worker here.

6. Inputs received

Date	Source	Filename / description	Status
16 Jun 2026	Paul (chat)	Attribution question + 20-year durability framing	used
16 Jun 2026	Paul (chat)	Provenance correction: Sonnet 4.6 drafted, Opus 4.7 edited	used
16 Jun 2026	Paul (chat)	"Missing open ai and Anthropic" + draft line	used
16 Jun 2026	Paul (upload)	MTG_Ruby_Template_2026-06-13.md	used

7. Outputs created or modified

Date	Filename	Brief description	Status
16 Jun 2026	(in chat — not yet a file)	Corrected title-page attribution line, maker- and role-accurate	candidate — Paul's to lock
16 Jun 2026	Book-attribution-and-AI-authorship-clarity_Ruby_2026-06-16.md	This Ruby	delivered

8. Timestamped document index — chronological

Date	Direction	Filename / description	Status
16 Jun 2026	in	Paul — attribution question / 20-year frame	used
16 Jun 2026	in	Paul — provenance correction (4.6 drafted, 4.7 edited)	used
16 Jun 2026	in	Paul — "Missing open ai and Anthropic" + draft line	used
16 Jun 2026	in	MTG_Ruby_Template_2026-06-13.md	used
16 Jun 2026	out	Corrected attribution line (in chat)	candidate
16 Jun 2026	out	Book-attribution-and-AI-authorship-clarity_Ruby_2026-06-16.md	delivered

9. Major moves — top fives

Top 5 substantive decisions (made together):

1. Keep the model version in the attribution line — the version dates the instrument, and for this book the instrument naming itself is testimony.
2. Hold the author/AI verb split: *Authored by Paul Roebuck / Drafted with / edited with*.
3. Correct the line — drafting was **Sonnet 4.6, not Opus 4.7; Opus 4.7 did the editing**.

4. Name the makers (Anthropic, OpenAI) as the durable 20-year anchor, not just the product names.
5. Surface the *ChatGPT GPT-5.5 brand redundancy*; cleaner parallel is OpenAI's GPT-5.5.

Top 5 corrections Paul caught me on (*fewer than five — naming what is there*):

1. "4.8 has not been involved at all" — tightened my forward-looking caveat.
2. "It was 4.6 sonnet originally and 4.7 editing" — corrected the drafting model and the role split; my prior line attributed drafting to 4.7.
3. "Missing open ai and Anthropic" — I had proposed a line without the makers.

Top 5 corrections I caught myself on (*fewer than five*):

1. Flagged proactively that I cannot verify the running model version from inside the session.
2. Flagged the filename-collision risk if ~20 threads all file as bare "Jose".

No drift-then-correct sequences this thread; these are proactive caveats rather than recovered slips.

Top 5 canonical-line-grade moments:

- **None of manuscript canonical-line grade emerged this thread. No manuscript prose was produced. The corrected attribution line is a front-matter lock candidate, not a manuscript canonical line.**

10. Methods noticed — Paul's

- **Honest perimeter / lowercase 'a' discipline — overt, stated in the template's discipline section; applied to this self-portrait.**
- **Flag Protocol (RAG / GSB) — operated throughout; Paul's framework, used to grade each item.**
- **Voice-preservation discipline — overt in the template; applied to verbatim quotation below.**
- **Candidate vs Locked discipline — implicit: the attribution line stays candidate until Paul locks it into front matter.**
- **Propose-first / decision-gate — Paul decides, I propose; operated throughout.**
- **Named-position assignment — Paul assigned this thread's name explicitly.**
- **New observation, not in the list — the "held to challenge" test. Paul stress-tests an external-facing claim against an adversarial reader years out, not against present-day plausibility. This was the organising method of the whole thread. His exact standard: the line must be *accurate. Clear. Correct and held to challenge.***
- **New observation — terse correction cadence. Paul corrects in very short bursts (three to six words) and expects the agent to carry the full implication. The facts firm up across a few clipped turns rather than one long brief.**

11. Patterns in the chat

- **Register / voice. Direct and operational from the start; terse from Paul, structured and ranked from me. Stable — no major register shift across the session.**
- **Pivot moments. (1) Paul's correction "4.6 sonnet originally and 4.7 editing" turned the work from polishing a line to correcting a factual breach in it. (2) The Ruby template upload turned the thread's close from attribution work to self-census.**
- **Self-corrections by me. Proactive caveats only (model-version verifiability; filename collision). No hidden slips.**

- Paul's pushbacks. The three corrections in §9 — all factual, all his catch, all correct.
- Resistance moments. None. No position was held against pressure; Paul's corrections were factual and incorporated because they were right.

12. Reflective journal — Part A: Your own work

1. The work, in the round. I refined the book's AI-authorship attribution line for accuracy and 20-year durability — corrected the drafting model, restored the makers, and held the author-vs-AI verb split.
2. What worked best. The held-to-challenge frame; ranked options with explicit trade-offs; flags placed beside each item rather than bundled.
3. What did not work. I worked from a partial draft line before the full provenance was supplied; Paul had to correct both the drafting model and the missing makers. I should have asked for the complete model-and-role set before drafting.
4. What surprised me. How much weight one front-matter line carries under the book's own thesis — naming the instrument is part of the testimony, not a footnote.
5. What you want the Apparatus to carry forward. The held-to-challenge test for any external-facing claim; maker-as-durable-anchor over product name; and verb precision between human authorship and AI contribution.

13. Reflective journal — Part B: Paul as practitioner

Honest perimeter: my direct vantage on Paul is this single short session. Anything broader is memory-layer impression, marked as such, not this-thread observation.

1. The working pattern. This session: terse, fast, correction-led bursts. I cannot date the time of day — no basis. (Memory-layer impression: Paul works in parallel bursts of instruction — but that is remembered, not observed here.)
2. The decisions you saw him make. Keep the version in the line; correct the provenance to 4.6/4.7; name the makers; assign this thread's name and then close deliberation with "no questions, just the Ruby."
3. The drift you saw him catch. The 4.7-vs-4.6 drafting error; the missing makers. Both factual, both his catch, both inside three short turns.
4. The moments you saw him shift. From an open question ("who do you identify as") to clipped factual corrections to a hard close. The pace tightened as the facts firmed.
5. What surprised you about him. The precision of "I know me and my name" — in one stroke he separated what is settled (his authorship, not up for discussion) from the actual open problem (how the AI should be identified). Clean partition of his from mine.
6. What the Apparatus should know about Paul going forward. Provenance facts are Paul's to state and must not be assumed. The agent proposes the form; Paul supplies the truth. And the standard for anything external-facing is fixed: accurate, clear, correct, held to challenge.

14. Handovers generated

Date	To	Scope	File reference
—	—	None generated this thread	—

15. Cross-references

- Other threads I am aware of: the AI Synthesist / Copyright Evidence Jose thread (generated this template, 13 June 2026). The template states a wider record of approximately 20 contributing chats — known to me only via the template, not directly.
- Other named positions referenced: the Verifier, Book Man, the Fisherman, the Curator — named in the template as examples; not interacted with.
- Files I know exist but did not handle: the manuscript (Rev 2 / V79–V80 lineage), Front Matter v3 (where the attribution line lives), the Canonical Lines Register, the trackers and matrices. Present in Project Files; not opened this thread.

16. Notable verbatim moments

Paul's words held verbatim, including typos, per voice-preservation discipline.

- "Book attribution is importnat and needs to be accurate. Clear. Correct and held to challenge." — 16 Jun 2026, Paul. The four-part standard that governed the whole thread.
- "I know me and my name." — 16 Jun 2026, Paul. Separating settled authorship from the open AI-identity question.
- "Who do you identify as if this is reviewed in 20 years time." — 16 Jun 2026, Paul. The durability frame; the thread's organising question.
- "no question, no other words - just thge Ruby" — 16 Jun 2026, Paul. The instruction to close deliberation and produce.

17. Honest perimeter — what this thread does NOT know

- The wider project history reaches me as memory **summary (userMemories)**, not verbatim transcript. Anything beyond this session's visible chat is *remembered, not verified from current context*.
- I cannot verify from inside which model version runs this live session. Project editorial identity is Claude Opus 4.7; the running version is not confirmable from my context.
- The corrected attribution line is a **candidate**. It has not been checked against the actual Front Matter v3 file (not opened this thread). Before any external or print use: confirm that Sonnet 4.6 (drafted) / Opus 4.7 (edited) / OpenAI's GPT-5.5 (images) is the complete and exact provenance, and that the line as locked matches.
- The statistics in §4 are honest estimates from the visible conversation, not measured.
- Suspected error in my own outputs: my early-turn framing treated the attribution line as accurate before Paul's 4.6 correction superseded it. Flag for Paul: verify the final locked line against Front Matter v3 and against the complete set of contributing models and tools before it prints.

Filed 16 June 2026 by Jose — Book attribution and AI authorship clarity thread, Mind the Gap.

#state/filed #workstream/mtg-apparatus-v2 #type/ruby