

Making AI work accountable.

Charles “Chuck” McFarland built an AI-assisted work system to help build Shed Happens, a wildlife data company that ships an app.

He brought construction-management discipline to AI: written work orders, evidence, separate inspection and owner approval. He directs the work, defines measurement rules and reviews labels; Claude and Codex help implement it.

Work with a record.

Each job carries its scope, tests, evidence and review into the next session. Production changes, external releases, new spending and public communications require specific founder approval.

Two examples, with receipts.

Approval stays specific. A recorded test found that one approval could clear two unrelated requests. The repair bound each decision to its own request. Saved fixtures show unrelated work remaining blocked.

A test can say no. In a new-scan test on 26 antler sides from 18 cards, a candidate increased independent tip matches from 51.2% to 63.4%, but produced more stray guesses. The improvement was uncertain, and the candidate failed its acceptance rule. It was not adopted.

October 27 status: an audit reported training/test overlap in the earlier model lineage. Chuck stopped the retrain and ordered a clean reset. Earlier training is treated as unproven; no validated new model is claimed.

Measure the resources, too.

A local-log report estimates \$12,633 of Claude API-equivalent usage in September and \$4,071 for October 1-27, against a stated \$200 monthly subscription. These are estimated usage values, not invoices, realized savings or proven productivity gains. Codex is separate.

The contribution is a practical way to bound AI work, inspect its claims and keep the evidence—including when the answer is “this did not pass.”

[Detailed case + four-page PDF / Evidence extracts](#)

Built with Claude and Codex. Prepared with AI assistance from dated project records. Private user locations and source records remain excluded.