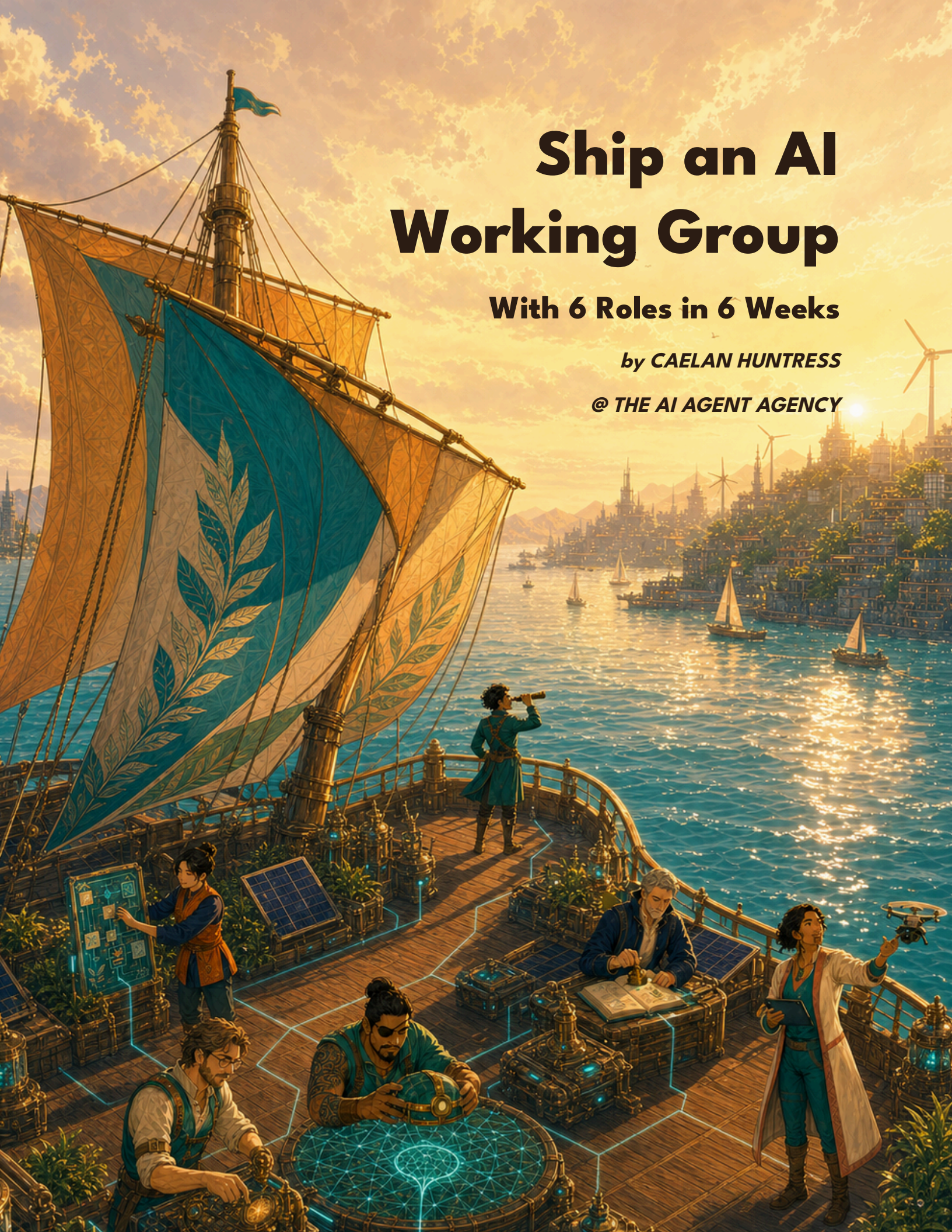


Ship an AI Working Group

With 6 Roles in 6 Weeks

by CAELAN HUNTRESS

@ THE AI AGENT AGENCY



Before we set sail...

COPY THIS THE RIGHT WAY.

You have permission to post, email, print, and pass this along to anyone you like — so long as you make no changes to its contents or format. Please do pass it along and make many copies. We reserve the right to bind it and sell it as a real book.

OUR CARE, YOUR RESPONSIBILITY.

This whitepaper is general in nature and is not a substitute for specific professional, legal, or technical advice. Please take specialist advice before acting on any idea here. The AI Agent Agency, GenAI Training LTD, and Caelan Huntress disclaim all liability to any person in respect of anything done in reliance, in whole or in part, on this document.

MANY SMART PEOPLE CONTRIBUTED.

The ideas here were shaped by working with real teams standing up real AI programmes — clients, operators, and practitioners who brought their troubles, skills, and hard-won lessons. The community is wiser than any one person in it, and this document is an accumulation of the wisdom and experiences of many.

A note on the research

The numbers in this paper are drawn from public 2025–26 studies (MIT, IBM, Gartner, ISACA, and industry reports). They are cited where they appear, so you can check the source and judge for yourself.



In 2026: Everyone is experimenting. Almost nobody is shipping.

The gap in AI roll-out right is *integration*. The tools are good enough. The missing piece is the way the company organises the work.

95%

of corporate AI pilots deliver no measurable return — stalled by the way the work is organised.

5%

integrate AI into how the company actually works. The right roles will build momentum.

Source: MIT NANDA, “The GenAI Divide: State of AI in Business 2025.”

When electricity replaced steam power, it took a generation to see the benefits. In a factory, replacing the steam boiler with an electric motor just turned the drive shaft a little more quickly. It was difficult to see a meaningful improvement until they tore up the factory floor, and redesigned the work around small electric motors.

To change the way the work gets done, you need time, and space, to focus on this change. That’s what an AI Working Group gives you - a crew to manage that transition. The companies that win with AI organise themselves to do more than just use the models well. They turn access into practice, practice into workflows, and workflows into measurable results.

In 2025, MIT’s NANDA initiative studied 300 public AI deployments and interviewed 150 leaders. The results were: 95% of corporate generative-AI pilots delivered no measurable return. Only 5% produced real value.

“The core barrier to scaling is not infrastructure, regulation, or talent. It is learning. Most GenAI systems do not retain feedback, adapt to context, or improve over time.”

Surveys in 2025–26 found that 98% of organisations have employees using unsanctioned AI tools, yet only 37% have any governance policy at all (IBM).

When something goes wrong, it is expensive: IBM put the cost premium on breaches involving “shadow AI” at over US\$650,000 *per incident*.

Why most AI Working Groups go nowhere

If a ship has no crew, it won't even get out of the harbour. When a crew gets on board a ship, there are two ways to fail, and most companies pick one accidentally.

1. **The Free-For-All.** No rules, no guidelines, no best practices. No momentum. Staff quietly paste contracts, customer data, and source code into whatever tool is open in another tab. It feels productive. It also lets data leak.
2. **The frozen committee.** The opposite failure is when a group is formed, and a recurring meeting takes up time in everyone's calendar. Maybe minutes are taken, and AI-generated reports are sent — but nothing ships. Everyone is so worried about doing AI wrong, they never do anything new at all. **The meeting becomes the work.**

Both failures share one root cause: **nobody owns any specific role.** Direction is just a guess. Projects have no deadline. Tools are unmanaged, and token spend is not planned or monitored. Knowledge evaporates between meetings. Nobody can tell the story of what got done.

The six roles in the following pages fix both of these failures at once. You need something simpler than a 40-page strategy, and more durable than another platform subscription: you need six people who each know exactly what questions to ask.

An AI Working Group needs six roles on the crew

A ship without a crew will drift, and never arrive; while a ship with a crew can sail to a destination. A committee can meet forever without progressing, while a crew has defined roles, develops rhythm, and knows how to harness momentum. Build your crew with six (or fewer) people - AI Champions can claim more than one role, if they wish.

The Captain	Sets and changes direction. Owns decisions, agenda, and priorities.
First Mate	Keeps projects moving to completion. Assigns project owners, monitors milestones, watches drift.
The Engineer	Maintains tech and tokens. Knows the models, moderates access, usage, budgets.
The Ledger	Maintains paperwork and audit trail. Policy, compliance, and AI Register.
The Skull	Builds the Company Brain. Captures and keeps institutional knowledge.
The Publisher	Communicates wins, inside and out. Sends the minutes, shares updates, makes case studies.

Assign roles. Meet weekly. Ship one low-risk workflow. Document what works. Share the result with the wider team.

Role 1: The **CAPTAIN**

(direction, priorities, time)

RESPONSIBILITIES

- Makes the final decision
- Schedules meetings & sets the agenda
- Keeps the group to time
- Sets priorities and changes course



Someone has to steer the ship. Without a single authority, a working group can debate forever, and commit to nothing. The Captain ends the frozen committee. Technical depth helps, and leadership, and management. But the core qualification is judgement: the ability to decide, prioritise, and protect the group's time.

MIT's research pointed at this directly: AI sticks under those leaders who can change real work. A distant central team can support an effort, but ownership has to sit close to the workflow.

Direction has to come from someone with the authority to change how the work actually gets done — and who has the nerve to kill a pet project, when it isn't working.

Questions To Ask

- Has the goal shifted since last week, or do we still move in the right direction?
- What's the biggest obstacle or challenge in our way?
- What decisions need to be made today?

Role 2: The **FIRST MATE**

(completion, owners, milestones)

RESPONSIBILITIES

- Catalogue every open project
- Archive stalled projects
- Assign an owner to each project
- Set milestones and check in on them



Ideas are cheap. Finished projects are valuable. The First Mate is the one who watches for drift — and is the person who makes sure that what gets started gets finished, or gets honestly buried.

The rule is simple and ruthless: every project has one owner, and a next milestone, or it goes in the archive. No orphan initiatives, no long list of ‘someday’ or ‘maybe’ projects. The First Mate maintains a short list of live projects, with names and dates.

Most AI projects stall quietly, waiting to be perfect. The First Mate pushes for a rough, shippable version that real people can use this month, then improves it from feedback.

Questions To Ask

- Which projects made progress this week, and which did not?
- What's stalled, and do we revive it or archive it right now?
- Does every live project have one owner and a next milestone?
- What can we ship this week, even if it's rough?

ROLE 3: The **ENGINEER**

(tools, access, tokens)

RESPONSIBILITIES

- Model selection and licences
- Access for team members
- Monitors token usage
- Sets token budgets



AI runs on two things: access and spend. Mismanage either, and you get one of two classic messes. Too little access and people route around you with personal accounts — that's how the average enterprise ends up with roughly a dozen AI tools, and IT aware of only a handful. Too little oversight, and the meter runs wild.

The Engineer chooses the right models for the job, holds the licences, grants and revokes access cleanly, and watches the token meter (so cost never becomes a surprise). They make the good path the easy path — so nobody needs a shadow tool to get their work done.

Questions To Ask

- Is anyone blocked on access to a tool or model right now?
- What is our current token spend, and could anything be a surprise?
- Is the model we're using actually the right one for the job, or just the default?
- Did anyone reach for an unapproved tool this week — and if so, why was the sanctioned path harder?

ROLE 4 - The **LEDGER**

(governance, audit trail)

RESPONSIBILITIES

- AI governance policy
- The AI Register
- Certifications and Credentials
- IT policy compliance and external compliance



In 2025 alone, more than 140 new AI-related laws were passed worldwide, while most organisations still operated with no governance policy at all. The gap between what teams are doing with AI, and what they can account for, is exactly where the risk lives.

The Ledger closes that gap. They write the policies, keep an AI Register of which tools touch what data, hold whatever certifications your industry needs, and maintain a record of use. They empower the group to answer, “Show me how you used AI here.”

Done well, this gives the group permission to move fast without fear. AI becomes an asset when you can prove what you did, why you did it, and who signed off when a client, an auditor, or a regulator asks.

Questions To Ask

- Is everything we shipped this week logged in the AI Register?
- Are we still compliant after this week's changes?
- Did anything touch sensitive data without a sign-off?
- If a client asked "how did you use AI here?" could we answer in 5 minutes?

Role 5 - The **SKULL**

(institutional knowledge)

RESPONSIBILITIES

- Protect the Company Brain
- Manage the team's access to it
- Ensure new knowledge is ingested
- Data structure, storage, and privacy



Every crew needs a flag. The Skull is yours — the keeper of the company's memory and AI identity. While the others sail the ship, the Skull makes sure the ship remembers everything it learns.

The Company Brain is a private, well-organised store of institutional knowledge that the team can query, according to their permissions levels. Get this right and your AI gets smarter every month, instead of starting from zero every session. MIT found the 5% who succeed share one trait: systems that learn & adapt to the company's workflows. The model is rented by everyone. Your edge is your knowledge: your playbooks, past projects, hard-won judgment — structured so AI can actually use it.

Questions To Ask

- What did we learn this week that the Company Brain doesn't yet know?
- What other bodies of knowledge can we access and incorporate?
- What knowledge would make next month's work faster or better?

Role 6 - The **PUBLISHER**

(internal & external comms)

RESPONSIBILITIES

- Weekly minutes for AI Working Group
- Production of Case Studies
- Monthly company updates
 - (via newsletter, video, or webinar)



Invisible progress withers. A working group that ships brilliant work, and tells no one, will quietly lose its budget, goodwill, and mandate. The Publisher makes sure that projects are turned into compelling stories, converting internal wins into tangible proof.

Inside the company, they broadcast what the group is doing — documenting the minutes each week, sharing a clear update each month — so the rest of the organisation sees momentum. People copy what they can see working.

The story of how you safely automated a workflow can also be a marketing asset, or an SOP. It can give your market visible proof that your AI capability is already working, and serve as training materials to help your company grow.

Questions To Ask

- Did anything ship that the rest of the company should hear about?
- Are any new wins complete enough to become a case study?
- Who outside this room needs to know what we decided?
- What's the one-line story of this week?

How the crew sails together

- The Captain **sets the course**
- The First Mate **drives projects**
- The Engineer **supplies the tools**
- The Ledger **keeps the documentation**
- The Skull **maintains the knowledge**
- The Publisher **broadcasts the results**

THE CADENCE

Weekly: a 45-minute AI Working Group meeting. Captain keeps it running on time, First Mate reports project status, Engineer flags spend and upgrades, Ledger notes anything to document, the Skull updates the Company Brain, the Publisher plans the communications.

Monthly: Publisher ships an update to the wider company. The Skull files what was learned. The Captain re-sets priorities.

Always: a human approves before anything customer-facing goes live. Clear gates are what make the crew fast. Everyone knows where human approval sits, so speed and trust can travel together.

“If you want to go fast, go alone. If you want to go far, go together.”

— AFRICAN PROVERB

Your first 6 Weeks with an AI Working Group

Start before the system is perfect. Name the crew, ship one thing, tell the story. You can refine everything else once the ship is moving.

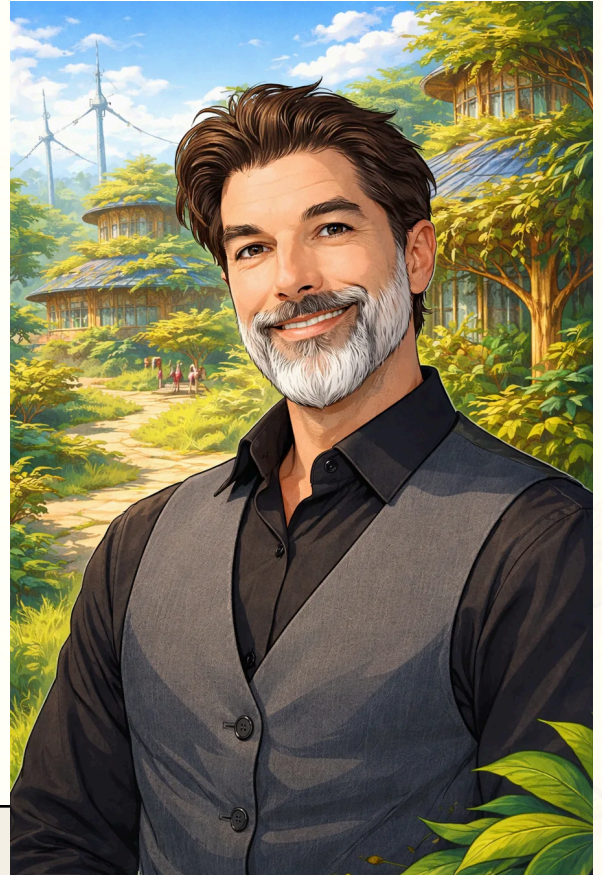
1. **Name a Captain.** One person who has the authority to make decisions.
2. **Assign the other five roles.** Double up where you must. A small team might be three people holding two hats each, that's fine.
3. **Hold the first meeting.** Take inventory: what AI work is already happening? (First Mate) What tools are already in use? (Engineer) What knowledge matters most? (Skull).
4. **Write a simple AI use policy.** The Ledger drafts it. Use a template. Define what's allowed, what data is off-limits, who to ask. Replace fear with clarity.
5. **Meeting 1: Pick a project to ship in 30 days.** Something small, useful, and low-risk. Make a workflow measurably faster. Resist the urge to boil the ocean.
6. **Meeting 6: Publish the first update.** The Publisher tells the company what shipped. Momentum is a story you keep telling.



Want a crew that ships? I'll help you build it.

I'm Caelan Huntress. I train teams across Australia and New Zealand to use AI for real work — practical, hands-on, focused on daily habits and the uniquely human parts of the job. I host the Christchurch AI Meetup, with more than 2000 members, and I have taught thousands of professionals to become confident operators of AI.

Through [The AI Agent Agency](#), I help small and mid-sized companies stand up an AI Working Group like the one in this paper, and deploy the agents that empower them to ship automatically.



How I can help

Stand up your six-role working group · Deploy practical AI agents and automations · Build your Company Brain · Train the crew to run the ship

I can play the First Mate, and facilitate your first 6 AI Working Group meetings, to give you the momentum to set sail on your own journey. If your AI effort is drifting — or hasn't left the dock — let's get a crew assembled and your next AI project shipped.

Book a call → caelanhuntress.com/schedule

Email → hello@caelanhuntress.com