Becoming Engineers

Six months of building AI frameworks and skills for knowledge work.

Barnaby Robson · June 2026

There are two kinds of people. Those who use AI to learn faster than they ever could, and those who use it to avoid learning at all.

— Mark Cuban

What this deck covers

A working AI stack for knowledge work, built over six months.

Eight levels, from chatbot to orchestrator

The higher the level, the more the system runs on its own.

1	**Chatbot**	ask, it answers
2	**Copilot**	works inside your files, beside you
3	**Agent**	runs a task step by step, asking approval
4	**Autopilot**	runs unattended, you review after
5	**Workflows**	a system that holds output to a standard
6	**Assistant**	works in the background, unprompted
7	**Multi-agent**	several long-running agents at once
8	**Orchestrator**	one agent runs a team of sub-agents

TODAY

Most people and organisations are here.

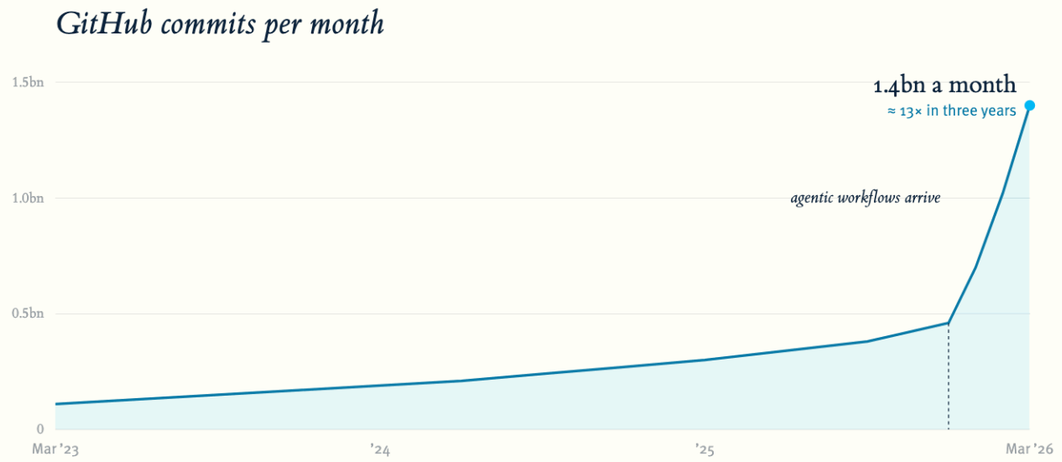
WHERE THIS PACK GOES

Reaching the **orchestrator** takes an engineering mindset. Each level up also widens the **blast radius**.

After the levels framing in Every's "An Executive's Guide to Implementing AI" (June 2026).

We are all going to have to become engineers

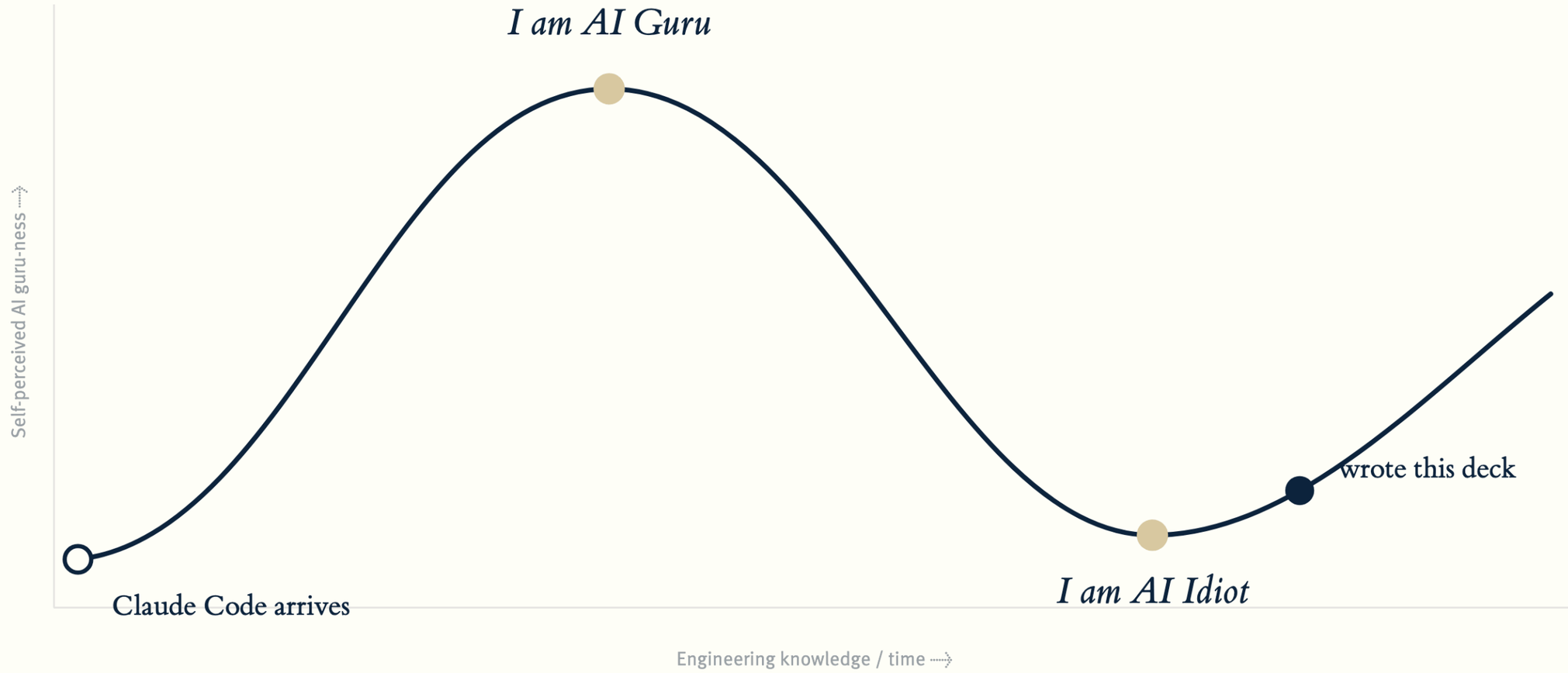
Nobody knows how AI reshapes work. The certainty: the skill worth learning is AI engineering — composing the skills, architecting the context an agent reasons over, building the harnesses. Underneath all three sits an older one: thinking in systems.



GitHub commits per month, March 2023 – March 2026. Source: GitHub, "An update on GitHub availability" (April 2026). Someone has to architect all of it.

Confidence peaks before the real learning starts

A weekend of wins makes you a guru; the broken week that follows is where the engineering education actually begins.



From the essay this deck precedes: “We are all going to have to become engineers”, barnabyrobson.org.

The code is just the shadow of the program. The program is the theory in the programmer's head.

— after Peter Naur, "Programming as Theory Building" (1985)

Agents now generate the shadow on demand — the theory of the system is still yours to hold.

The skill underneath it all is systems thinking

A system is the pattern of how the parts affect each other over time — change one and the others react; miss a connection and the pattern breaks. Senior engineers built this sense by accident, over years of failure. Working with agents, you need it on day one.

The orchestra	AI plays any instrument on demand — violin, drums, brass — often better than a human can.	**the model**
The conductor	Someone still decides how the parts fit together — when the strings hold back, when the brass comes in.	**you**

After the orchestra framing in Machinations, “Is this the only skill left?” (April 2026).

Three questions, before anything runs

Systems thinking made concrete. Each one is answerable from the architecture alone — before a single step executes.

01 · STATE

Where does state live?

Who owns the truth?

If two parts of the system each believe they own it, there is a bug already. It just hasn't been triggered yet.

02 · FEEDBACK

Where does feedback live?

How do you know it's working?

Logs, metrics and errors, bubbling up somewhere you look. A system that tells you nothing is pretending to work.

03 · DELETION

What breaks if I delete this?

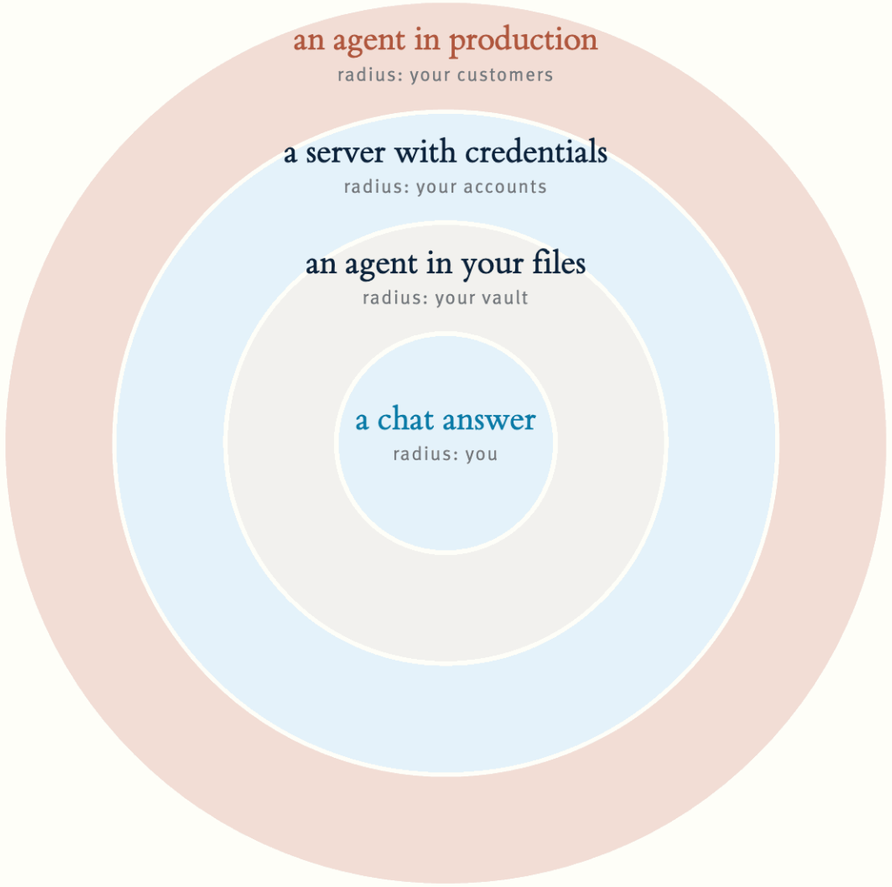
Can you trace it in your head?

Pick any component and trace its blast radius before touching it. When the answer is "I don't know", that is the homework list.

The three questions follow Hak (Machinations), "Is this the only skill left?" (April 2026).

Blast radius — how far a mistake travels

Once an agent can act, the hard problem moves from intelligence to control. Before granting a new power, know the radius of its worst day — then shrink it.



A coding agent reportedly deleted a live AWS production environment — a 13-hour outage. A Cursor agent wiped a company database, and its backups, in nine seconds.

allowlists	the agent touches only what is named
confirmation gates	destructive steps wait for a yes
narrow scopes	one job, one folder, one credential
crawl-walk-run	new powers arrive in stages, watched

PC Gamer on the AWS/Kiro report (Amazon attributes the outage to misconfigured access controls); Tom's Hardware on Cursor/PocketOS. Each level up the eight-level ladder widens the radius — that is what the engineering is for.

Every layer below you used to be verifiable

Assembly to C to Python: each abstraction translated deterministically, so you could trust the layer without understanding it. The new layer is different in kind.

Compiler	Same input, same machine code, every time — **provably correct**, a guarantee underneath you.	**trust without understanding**
Interpreter	The same contract — nobody reads the bytecode, because nobody needs to.	**trust without understanding**
LLM	Same input, **different output each run**. It can introduce a security hole or a wrong business rule without telling you.	**trust earned by checking**

The reason the harness loop ends in a Check, and every deliverable in this stack passes a review gate before it ships.

Split the work: deterministic and stochastic

After enough broken decks, two old words re-enter the vocabulary. Sort every job into the half that must repeat exactly and the half that improves when it varies.

DETERMINISTIC · SCRIPTS OWN THE RULES

The same result, every time

- Brand palette, slide master, font sizes, chart geometry.
- Cross-foots, totals, the underlying financial analysis.
- They belong in **tested scripts** the model runs — obeyed every time.

STOCHASTIC · THE MODEL OWNS THE FILL

Better when it varies

- Narrative, framing, emphasis, the visual idea.
- Creative work — what the models are genuinely good at.
- Give the model the frame and let it **invent inside it**.

Ask one prompt to do both and it quietly breaks the deterministic half — open the file and the logo sits two centimetres into the bleed.

The skill format later in this section embodies the split: method in markdown, rules in scripts/.

Build modular, so any part can swap

The frontier moves weekly — models leapfrog, token prices lurch, a vendor can close a door overnight. Architecture that survives this is made of replaceable parts.

Models

Tier the work and keep a **tested fallback**. When one provider shut OAuth for agents one April morning, the workflows moved — and the knowledge graphs came along intact.

Knowledge

Plain markdown with metadata, owned on disk. Apps die and platforms pivot; **files outlive both**, and any model can read them.

Skills

Forty loose skills is a hammer collection. **Families with routers** — one per domain, named workflows inside — keep instructions from competing.

Munger's warning rings here: if all you have is a hammer, everything looks like a nail. The six families later in this deck are the principle applied.

The silos collapse into one operator

Agents supply the depth in any lane, on demand. The scarce seat holds the whole picture and takes the judgement calls.

Software	Backend, frontend, devops, database — one builder now ships across all of them, with agents carrying the lane depth.	**one engineer**
Knowledge work	Analyst, designer, editor, publisher — the same collapse. One person, one stack, every phase of the work.	**one operator**

Everything from here is the demonstration: one person running capture, organise, process and execute.

Three years, three engineering disciplines

One operator conducting everything needs a craft, and the craft has climbed a layer each year. The next three slides take the layers in turn.

2023

Prompt engineering

owns the message

What you say to the model. Picked up in an afternoon — the easy layer, and getting easier.

2024

Context engineering

owns the window

What the model has read when it answers. What loads, in what order, with what anchoring.

2026

Harness engineering

owns the loop

The system the model runs in: rules, fresh context per task, tools, and a gate before shipping.

Prompting can be picked up in an afternoon; the other two layers are the engineering.

A skill is a folder the agent loads when it is needed

Prompt engineering, industrialised. Write the method once, and every agent runs it the same way — the unit of reuse.

```
bvest-analyst/
├─ SKILL.md          ◀ the instructions — the working method, steps and standards, in plain markdown
├─ scripts/          ◀ the deterministic parts — a tested script the model runs where improvising would slip
│  └─ screen.py
└─ references/       ◀ deep knowledge, pulled into context only when the work calls for it
   └─ koch-star.md
```

```
description: "use when screening a stock — quick fundamentals and a tear sheet"
```

▲ one line decides when the skill loads, so the context window stays clear until it is needed

Sixty skills in this stack, grouped into six families. Anthropic shipped the format as Agent Skills in late 2025; the stack has run on it since.

The context window is the constraint

Context engineering is what the model has read when it answers. A model forgets at the session boundary and repeats mistakes across them; the plumbing around the window carries memory in, hands work across before it fills, and makes each session teach the next.

recall

At the start, load what was in flight from the last session — by date or by topic.

handoff

Before the window fills, package the live state and pass it to a fresh agent, keeping detail a bare summary would lose. Nothing in progress is lost.

retrospective

At the end, work out what was learned and write it back into the skill.

Three of the operating skills. skill-manager and grill-me round out the kit underneath. Recall, handoff and retrospective answer the first systems question: where does state live.

A harness is the loop that turns a model into an agent

A model on its own answers once. A harness wraps it in a loop — tools, an action, the result fed back — running again until the work is done. This is harness engineering, the third layer of the craft.

01
Plan
Read the goal and the context, decide the next step.

MODEL

→

02
Act
Call a tool — read a file, run a script, search, edit.

TOOLS

→

03
Observe
Feed the result back in — the error, the output, the diff.

LOOP

→

04
Check
Does it meet the goal? If not, go again.

GATE

repeat until it holds, then stop

`/loop` wraps any prompt in this loop — on a clock, or self-paced. **The harness, as one line.**

`autoresearch` `autoloop` the same loop with a **verify-and-keep gate** — iterate toward a goal, keep only what improves.

Claude Code, Codex and cmux are harnesses. The skill supplies the method; the harness supplies the loop, the tools and the stopping rule. The Check gate answers the second systems question: where does feedback live.

The big prompt gives way to the coordinated loop

The payoff of the harness: decompose the goal, give each job the right role, gate the result — and stop being the bottleneck.

BIG PROMPT · ONE AGENT, EVERYTHING AT ONCE

Re-prompt and pray

- ✕ “Do all of this: research, build, review, ship, and don’t forget the edge cases…”
- ✕ One agent juggles every role and hands back **one answer**.
- ✕ Guesses ship — nothing stands between the answer and you.

COORDINATED LOOP · THE RIGHT ROLE FOR EACH JOB

Decompose, assign, gate

- The goal is **decomposed**: research, execution and review each get their own agent.
- Output is reviewed **against the original goal**; a failed gate goes back with a reason.
- What worked is **saved**, and you step in only at decision points.

After @jumperz on X. Level 8 on the ladder — the orchestrator — is this loop, running.

The stack reads markdown and ships HTML

A person and a model both read **markdown** without a parser. The finished artefact renders as **HTML** — one self-contained file the model writes itself, styled inline, clean in any browser.

POWERPOINT & WORD

Binary, and in the way

- A container a model **cannot read** without a converter.
- Layout welded to content; an edit means fighting the tool.
- Heavy files and version sprawl, always a render step from the web.

MARKDOWN & HTML

Plain text in, HTML out

- Markdown in: a model parses it **natively**, a human reads it raw.
- HTML out: **one file**, styled inline, opens anywhere, prints clean.
- The model writes the finished artefact itself — no export step.

This deck is a single HTML file, written in one sitting.

The investment tear sheets, the published data-stories and this pack all render as HTML from markdown; client deliverables still export to Office when the engagement needs them.

The fundamentals predate the models

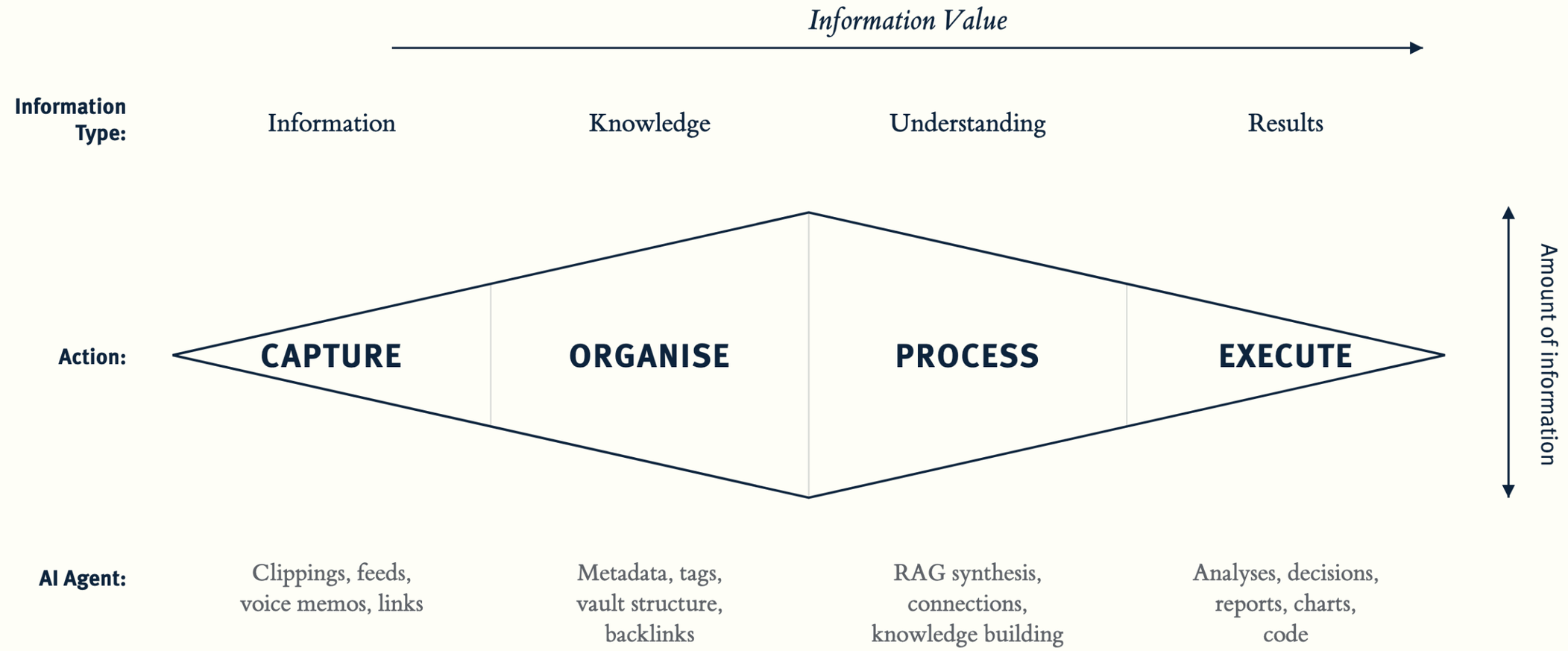
Dave Plummer wrote Task Manager and spent the NT years inside Windows. His advice to engineers, mapped to the agent era.

“90% of the job is debugging”	Most agent work is reading logs and fixing workflows — the dip on page six is the curriculum.
“Manage complexity from day one”	Slop compounds and becomes unfixable. Build small, crawl-walk-run, prune the rules weekly.
“Tools and processes matter more than you think”	Good tooling amplifies your output. The harness, the scripts and the review gates are the investment.
“Understand the problem before writing”	Design before you prompt: draw the system, write the spec, then let the agent build.

Dave W Plummer on X. His fifth lesson — people and politics beat pure tech arguments — is the enterprise story, kept for another deck.

COPE: capture, organise, process, execute

GTD and PARA, rebuilt for agents. The model does the bulk of each phase; you set the frame and take the call.



Putting COPE to work

Everything from here is my own AI stack — the b-stack — COPE put to work across capture, organise, process and execute, built over six months. The principles behind it, and the skills themselves, are reusable: the point is that you can build your own.

Six skill families, each owning a phase

Past a certain count, loose skills become a heap of competing instructions. Families fix that — **sixty skills and about sixty workflows, in six**. Capture runs on rails, so it needs none.

CAPTURE

on rails

clippers & feeds

web, email, voice ⇢ markdown

ORGANISE

into knowledge

bknow

frontmatter, backlinks, synthesis, integrity

10 skills · 7 workflows

PROCESS

the reasoning

bdas

the deal lifecycle

10 skills · 7 workflows

bvest

the investment decision

11 skills · 42 workflows

EXECUTE

to a standard

blook

charts, decks, databooks, on-brand

10 skills · 3 workflows

bvoice

writing in one voice

12 skills

brun runs underneath all four (7 skills) — health checks, cron and repair keep the fleet alive.

Sixty skills, six families, built in six months

Roughly 3,600 agent sessions on Claude Code and Codex. The engineering that came out — harnesses, extractors, renderers and review gates. A sample of the stack:

CAPTURE
on rails

Markdown ingester
PDF, OneNote and email ⇢ clean markdown, on-device

Headless VDR extractor
drives real Chrome over CDP through a locked data room ⇢ markdown; 930 files

Web clipper
any page ⇢ clean markdown, filed in the vault

Podcast transcriber
a YouTube episode ⇢ transcript, quotes and highlights

ORGANISE
into knowledge

Semantic vault search
BM25 + vector + rerank over ten vaults, 94k links

Karpathy wiki compiler
raw, wiki, schema — the model owns the synthesis

Clippings processor
categorise, topic, backlink and file each clip

Vault integrity
frontmatter, wikilinks and filing held to one standard

Person CRM
people notes kept current from meetings

PROCESS
the reasoning

bdas

Databook reconciler
cross-foots every column and ties each figure to source; 39 sheets

Proposal engine
an RFP to thesis, scope, fees, storyboard and deck

Commercial DD engine
multi-agent industry research and expert-source synthesis

bvest

Five-persona committee
Mauboussin, Munger, Druck, Koch, Perruna vote a package

Deep-dive extractor
a fiscal.ai databook ⇢ a 20-metric value-creation package

Macro synthesiser
capital-cycle signals ⇢ a seven-factor conviction score

xlsx valuation model
a databook ⇢ DCF, multiples and a scenario tree

Investor-wiki builder
153 Mauboussin papers; 183 Munger mental models

EXECUTE
to a standard

blook

Data Insights
the blook engine and Tufte principles — on-brand analytics; 447 MPF funds shipped

HTML slide decks
a deck-spec ⇢ a self-contained HTML deck and PDF (this pack)

KPMG deliverables
report, deck, note and databook in the firm brand

Figma studio
brand tokens ⇢ on-brand components, over MCP

bvoice

Voice pipeline
storyboard ⇢ first draft ⇢ review, one voice

Editorial gate
a six-lens, write-time quality check

Wait — how much did we spend?

Six months and roughly 3,600 agent sessions carry a real bill. Even the teams shipping fastest are recalculating — budgeting the spend is part of the craft.

APPLIED AI

Uber CTO shows how Claude Code can blow up AI budgets

Uber's surging use of AI coding tools, **Anthropic's Claude Code** in particular, has maxed out its full-year AI budget just a few months into 2026, according to chief technology officer Praveen Neppalli Naga.

"I'm back to the drawing board because the budget I thought I would need is blown away already," Neppalli Naga said in an interview.

Uber is one of a number of companies now running internal leaderboards ranking the software engineers who use the most AI.

The Information · 2026

Capture

Everything in, as markdown. The one phase with no skill family — clippers, MCP tools, converters and cron jobs do the work as content lands.

Everything becomes markdown

A model reads markdown straight, no parser. So every intake lane ends the same way — clean markdown in the vault inbox, whatever the source.

Tagged [[Inbox]] as it lands — waiting for Organise to file it.

A server that captures while I sleep

Hermes, a headless agent on a Hetzner box, runs the scheduled lanes around the clock — so the morning starts with the vault already filled.

Scout lanes	**Idea, theme and anchor** — market views captured as they appear.	**bvest vault**
Newsletters & podcasts	Transcribed overnight and parsed for themes.	**bvest vault**
Gmail to vault retrieval	The inbox routed to the right vault note.	**DAS vault**
Clippings sweep	Web clips processed and filed on a schedule.	**Core vault**
Morning brief, evening roundup	The day's signals, written into the daily note.	**daily note**

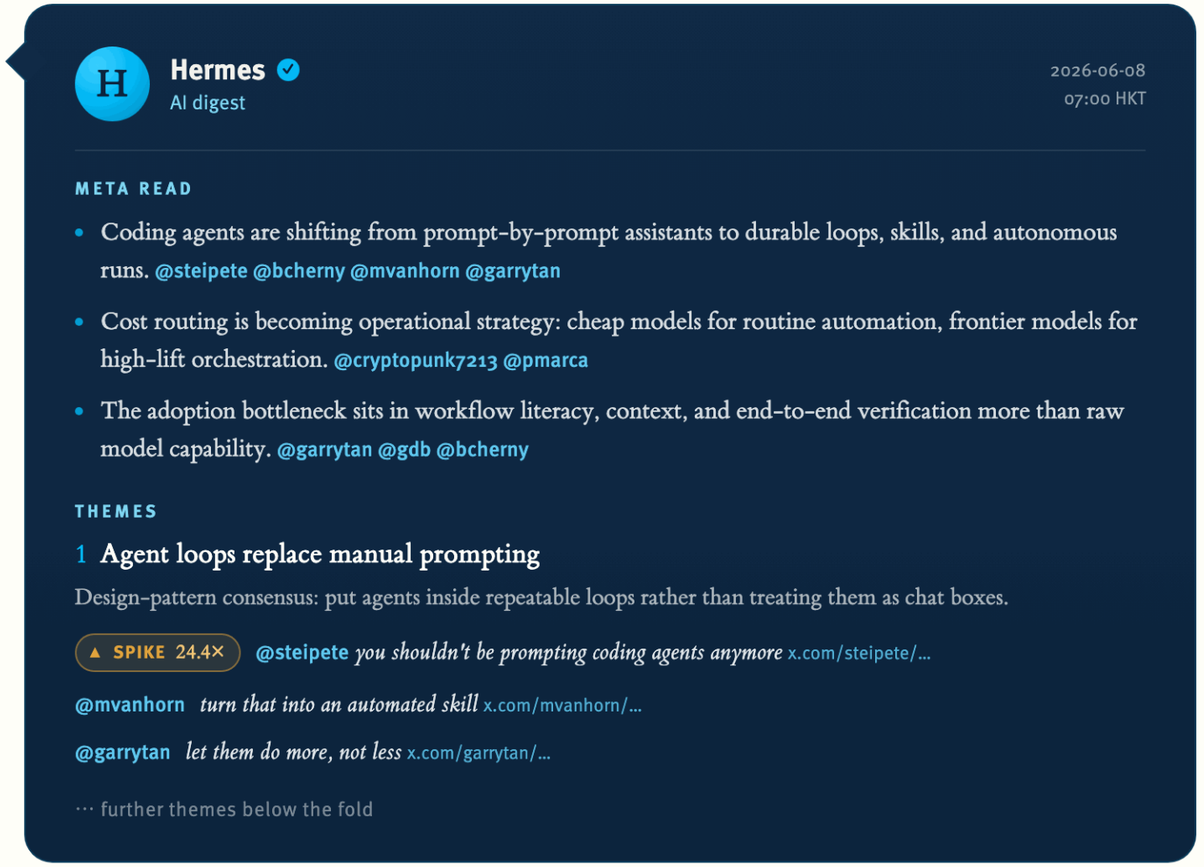
Hermes absorbed the old OpenClaw jobs in June 2026 — one server now owns every scheduled capture lane.

The 07:00 digest, as it lands

A real Hermes AI digest, recreated — the overnight X sweep compressed into a meta read and ranked themes, every claim carrying its handles and links.

The point is the compression. A night of posts becomes three meta-reads and a handful of themes — and a 24.4× engagement spike is flagged before breakfast.

overnight X sweep ⋯⟩ Telegram · 07:00 HKT



Recreated from the live Telegram message of 2026-06-08 — meta read, themes and spike flags as delivered.

Organise

*Into knowledge. **bknow** files what Capture lands — frontmatter, backlinks, integrity — and compiles it into a wiki an agent can read.*

Every clip becomes a structured, linked note

It lands tagged **[[Inbox]]**. **bknow** reads the content, writes the frontmatter, files it to the right folder and links it into the graph — first mention, whole word, longest match — all held to one vault standard.

Methodologies/FDD/Quality of Earnings.md

```
categories: [[Methodologies]]
topics: [[FDD]], [[EBITDA bridge]]
service_line: [[Due Diligence]]
source: [[FDD Training Deck.pdf]]
created: 2026-06-09
status: active
```

Quality of Earnings

Adjusts reported EBITDA for one-off, non-recurring and owner items, so the [[EBITDA bridge]] carries through to sustainable earnings. It sits at the centre of [[Financial Due Diligence]] — the number every [[Valuation]] multiple is applied to.

- **Frontmatter** written to schema
- **Topics** tagged
- **Source** linked back
- Filed to **Methodologies/FDD/**
- **4 backlinks** wired into the graph

One command brings every vault current

bknow-sweep walks every vault under the root, reads each one's type from its CLAUDE.md, and dispatches the right maintenance agents in parallel — refresh search, process inboxes, validate links, update cross-references — then writes one consolidated report.

Agents/Audits/sweep-2026-06-10.md CONSOLIDATED

7	6,832	6	1
VAULTS DISCOVERED	DOCUMENTS INDEXED	AGENTS IN PARALLEL	DETERMINISTIC PASS

VAULT	TYPE	INDEX	INBOX	CLIPPINGS	LINKS	WORKSTREAMS
Core	personal	4,338 ✓	6	3	1,204 ✓	—
Project X	engagement	1,052 ✓	1	—	642 ✓	5
Project Y	engagement	749 ✓	0	—	588 ✓	7
DAS	practice	418 ✓	2	1	311 ✓	4
bvest	practice	141 ✓	0	—	96 ✓	2 ⚑
bdas wiki	wiki	134 ✓	—	—	128 ✓	—
Agents	operational	live	—	—	410 ✓	—
All vaults		6,832	9	4	3,379	18

source: /bknow-sweep · vaults walked from ~/Obsidian/*/CLAUDE.md, each type read and dispatched in parallel · QMD re-embedded, 0 vectors pending · 4 broken links repaired, 2 orphans flagged · ⚑ one cross-ref held pending its source meeting · full QMD store 8,206 docs incl. five persona wikis

The same fan-out pattern as a dynamic workflow — many vaults, isolated agents, one deterministic pass.

An agent is only as good as what it has read

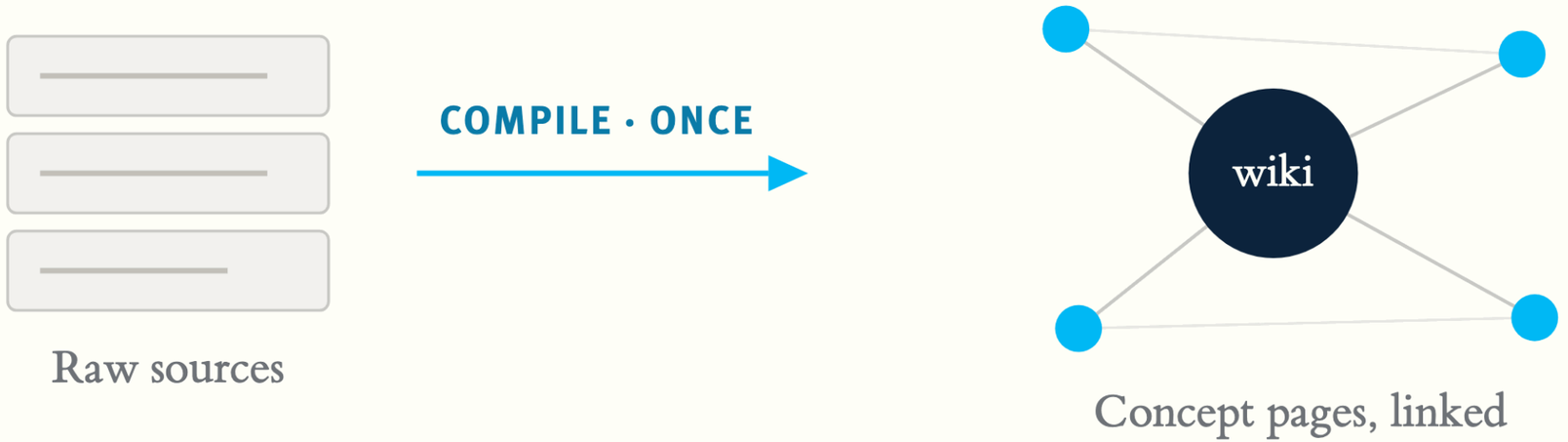
Behind each process family sits a vault built on Karpathy's three-layer wiki. The agent reasons over a compiled body of work, retrieved on demand. In his framing, Obsidian is the IDE, the model is the programmer, and the wiki is the codebase.

Raw	**The source.** Papers, methodologies and books, enriched to markdown. The curator drops them in; the model reads them and never edits.	CURATOR OWNS
Wiki	**The synthesis.** One page per concept — summary, mechanism, application, contradictions flagged, provenance carried.	MODEL OWNS
Schema	**The rules.** How pages are structured and which workflows run — ingest, query, lint — the file where the two co-evolve.	CO-EVOLVED

After Andrej Karpathy's LLM-wiki pattern (April 2026): a knowledge store compiled for a model to read.

Compile the sources once

The reflex is a vector store, returning fresh chunks on every query. Karpathy's reframing is a compiler analogy: the curated wiki pays the synthesis tax a single time — concept pages, hubs and **[[wikilinks]]** the model keeps current — and every query after rides that structure.



Pay the synthesis tax once at ingest, then keep the artefact current — each ingest revises the pages, each good answer is filed back.

The reusable wiki, and the one built for a single deal

Both are Karpathy wikis, differing in lifespan and in what they hold. One compiles methodology and mental models, cited across every engagement. The other is the living record of one deal. The first feeds the second.

KNOWLEDGE WIKI

Reusable, permanent

- Methodology and mental models — **9 Levers of Value**, Quality of Earnings, ROIC, the latticework.
- Raw sources, compiled into concept pages, governed by a schema.
- Cross-engagement: written once, cited on every deal.

e.g. **DAS** (420 notes) · **Mauboussin** (377 pages, 13.9k links)

PROJECT WIKI

One deal, time-bound

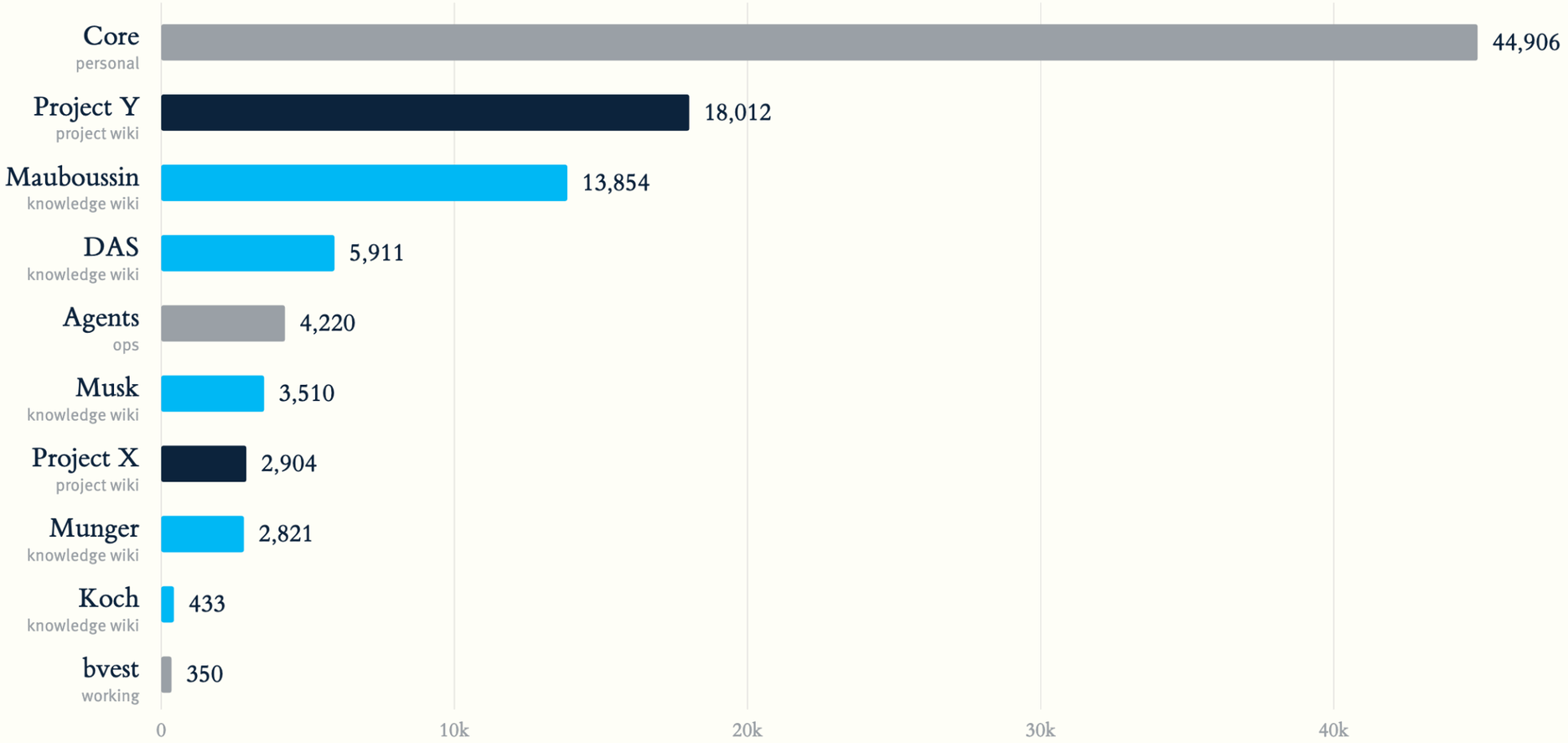
- The engagement's living record — **people, meetings, the data room, workstreams, the databook**.
- One note per file, one page per person: the deal's CRM and source of truth.
- Stood up at kick-off, archived at close.

e.g. **Project Y** (752 notes, 190 meetings, 223 people, 18k links) · **Project X**

The knowledge wiki feeds the project wiki — methodology applied to a live engagement.

13,500 notes, 97,000 wikilinks

Every note carries its date, source, topic and links, so the vault is a graph the model can traverse — integrity-checked and orphan-swept by bknow. The knowledge and project wikis are its densest neighbourhoods.



Wikilinks per vault, ten vaults. The two project wikis diverge — Project Y's meeting graph runs to 18k links, while Project X stays note-heavy (1,054 notes, 930 of them VDR index) and lighter on links. Queried via the graph view, Bases, backlinks and semantic search.

The deal

A buy-side diligence across the full lifecycle — go-to-market, fieldwork, analysis, reporting — each phase with its own skills, all feeding one project vault.

Four phases, pitch to report

One engagement, four phases — each with its own skills, all feeding one project vault the model reads back.

01

GTM

Frame the thesis and win the work — research, the value-driver tree, the 9 Levers, the proposal.

FRAME & WIN

02

Fieldwork

Gather the evidence — the VDR extractor and management meetings, wikilinked into the vault.

GATHER

03

Analysis

Build the numbers — captures reconcile into a cross-footed databook.

ANALYSE

04

Reporting

Tell the story — findings compile into the KPMG report, on brand.

REPORT

Every phase reads one methodology vault

90 methodology notes, compiled into 33 concept pages. bdas queries it semantically — the concept first, then the source note behind it.

9 Levers of Value	VALUE CREATION
Quality of Earnings	FINANCIAL DD
Net Debt & Debt-Like Items	COMPLETION
Completion Mechanics	SPA
Commercial Due Diligence	STRATEGY
Synergy Assurance	INTEGRATION
TSA Design	SEPARATION
Target Operating Model	VALUE CREATION

DAS METHODOLOGY VAULT

33

concept pages, compiled from **90** source notes

Retrieved live via semantic search, across five service-line hubs — the same Raw / Wiki / Schema layering as every vault.

Where value sits, before the work starts

Before any fieldwork, the thesis is framed — what the market is, and where value is made and lost.

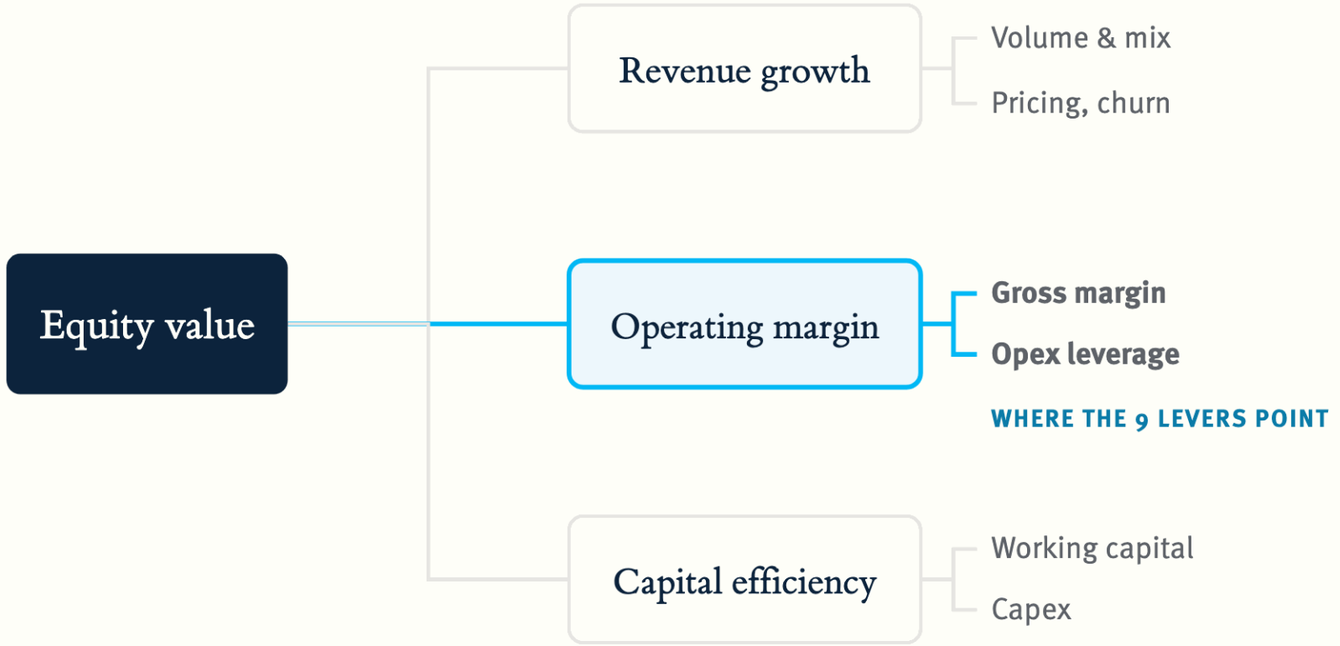
BEFORE THE WORK STARTS

Frame where value sits

Read the market — the commercial-DD engine runs multi-agent web research, each claim cited.

Map the value — the value-driver tree decomposes equity value, and the 9 Levers point the diligence at the few drivers that move the deal.

Both feed straight into the proposal.



A separation proposal, built over one weekend

Project Ivory — a cross-border retail-banking separation. The RFP arrived; over a single weekend in Claude Code, the agent produced the full response. Every page here is rebuilt from that response, **anonymised**.

Hard close readiness — Day 1 with no safety net

The deal terms rule out transitional support after Day 1, so the response maps every traditional TSA protection to a pre-close readiness equivalent.

TRADITIONAL TSA SCOPE	HARD CLOSE READINESS EQUIVALENT
Schedule design	Separation completeness checklist — every dependency tracked to closure
Service delivery	Pre-close collaborative arrangements — SME support, data access regimes
Exit management	Buyer readiness tracking — tollgate evidence at each milestone
Post-close stabilisation	Pre-close dress rehearsals, three per close date — problems surfaced before Day 1
Pricing & SLA governance	Data dictionaries, product mapping, structured face-to-face sessions

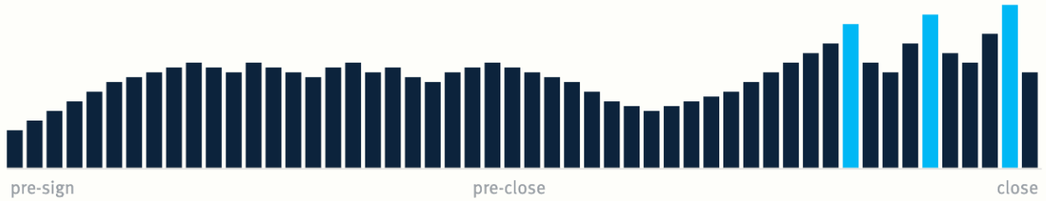
The steering committee's Go/No-Go evidence base, with every protection rebuilt as pre-close evidence.

The methodology page, rebuilt

Resource summary

Fixed-team cover for advisory, PMO and readiness scopes, phased weekly across the programme.

1,518	905	200	413	~13m
TOTAL PERSON-DAYS	CORE ADVISORY & PMO	WORKSTREAM SUPPORT	HARD CLOSE READINESS	PROGRAMME DURATION



Weekly person-days across the programme — the late spikes are the three dress rehearsals.

The resourcing page, rebuilt

Deal intelligence, a 25,000-word precedent study, scope and risk register, resourcing and the fee model, storyboard and pitch deck — RFP in, a full response out for the Monday.

Meetings become structured notes, on their own

Half the deal is what people say across months of meetings. Each one is transcribed; the agent turns the transcript into a structured, wikilinked note.

Project X/Meetings/2026-06-09 Solvency & Reserving Walkthrough.md

```
type: [[Meetings]]
date: 2026-06-09
project: Project X
people:
  - [[Target — CFO]]
  - [[Target — Appointed Actuary]]
  - [[KPMG — Engagement Leader]]
topics:
  - [[Solvency]]
  - [[RBC ratio]]
  - [[IFRS 17]]
workstream: [[Financial DD]]
created: 2026-06-10
```

Project X — Solvency & Reserving Walkthrough

INFO TO PROVIDE

- **Capital position**. The RBC ratio rose across the diligence period — 259% at FY23, 293% at FY24, 326% at Sep'25 — on available capital of HK$3,049m against required HK$935m.
- **Reporting basis**. The Sep'25 figures sit on the IFRS 17 transition; third-quarter technical reserves remain in draft pending the Q3 actuarial pack, due within two weeks.

KEY INSIGHTS

- **The solvency build is real but front-loaded by one reserving release**. Roughly a third of the FY24 uplift came from a single prior-year release on the motor book; the confirmatory work should isolate the recurring movement.
- **Reserve adequacy gates the headline**. Until the Q3 pack reconciles to the audited technical reserves, the Sep'25 figure carries one flagged input. Treat it as indicative until the pack lands.

ACTIONS

#	OWNER	ACTION	DUE
1	**Target (Appointed Actuary)**	Release the Q3 actuarial pack and reconcile technical reserves to the audited Sep'25 position.	w/c 2026-06-23
2	**KPMG (Engagement Leader)**	Isolate the recurring solvency movement from one-off reserve releases in the databook.	Pre-confirmatory

Note recreated from a Project X deal meeting; names and figures anonymised. At scale on Project Y, the six-month engagement: **190** meeting notes · **223** people pages · **28,465** lines of structured minutes, every name a wikilink.

Getting data out of a locked room

A data room is built to stop exactly this — IRM-locked, downloads wrapped in DRM, copy disabled. The VDR skill drives my own Chrome through the secure viewer and indexes the whole room. Built and hardened on Project X, a live buy-side diligence on an insurer.

CAPTURE

Read the room

- Playwright attaches to my **real Chrome over CDP** — a headless browser gets blanked by bot detection. I do the login and MFA by hand.
- Each document opens in the secure viewer. Read the spreadsheet cells **straight off the page** for exact values; screenshot the tables that resist.

ORGANISE

Build the VDR base

- Crawl the folder tree into **one note per file**, mirroring the room's own numbering.
- An Obsidian Base over those notes becomes the **live extraction queue** — everything, by area, done, blocked, by workstream.

PROJECT X

930

files crawled into the VDR base — the whole room mapped and queued

442

IFRS-17 captures from one section alone; entity names DRM-redacted at source

Captures become a reconciled databook

Raw captures transcribe into a reconciled model — cross-footed, classified by deal impact, every figure traceable to its source document.

RECONCILE ⋯⋯ ANALYSE

Cross-foot, then classify

One sheet per source tab, units and reference in the header. **Cross-foot every column**, reconcile row counts against the room, spot-check ten rows. Uncertain cells flagged.

Then run the clean numbers against the methodology vault — quality of earnings, net debt, the 9 levers.

39 databook sheets on Project X, each cross-footed and reconciled.

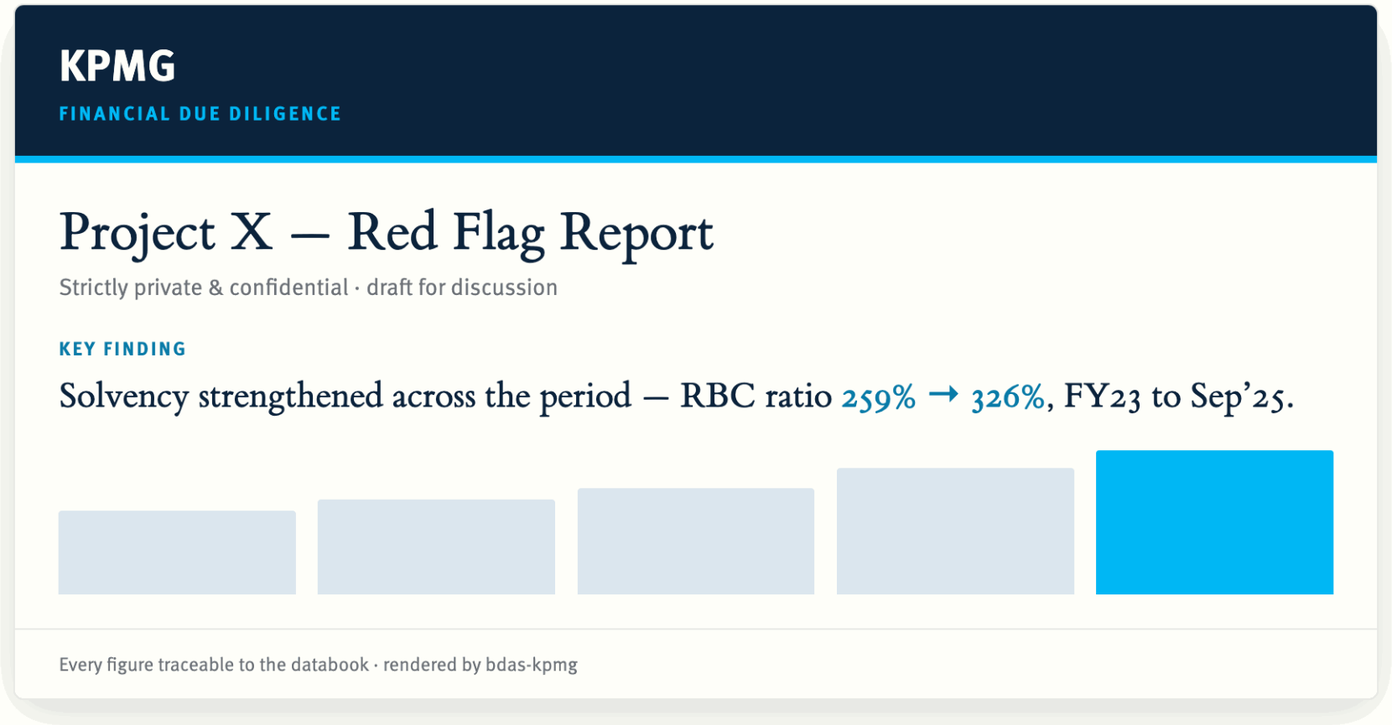
Project X — Solvency.xlsx RECONCILED

METRIC (HK$M)	FY23	FY24	SEP'25	
Available capital	2,418	2,706	3,049	✓
Required capital	933	924	935	✓
RBC ratio	259%	293%	326%	✓
Net earned premium	1,476	1,612	884	✓
Technical reserves	4,102	4,380	⚑	⚑

source: Audited financial statements FY23–Sep'25 · tab 4.2 — one cell flagged, awaiting the Q3 pack

Findings become the deliverable, on brand

bdas-kpmg owns the KPMG brand contract — palette, fonts, provenance — so every output renders on brand, every figure traceable to the databook.

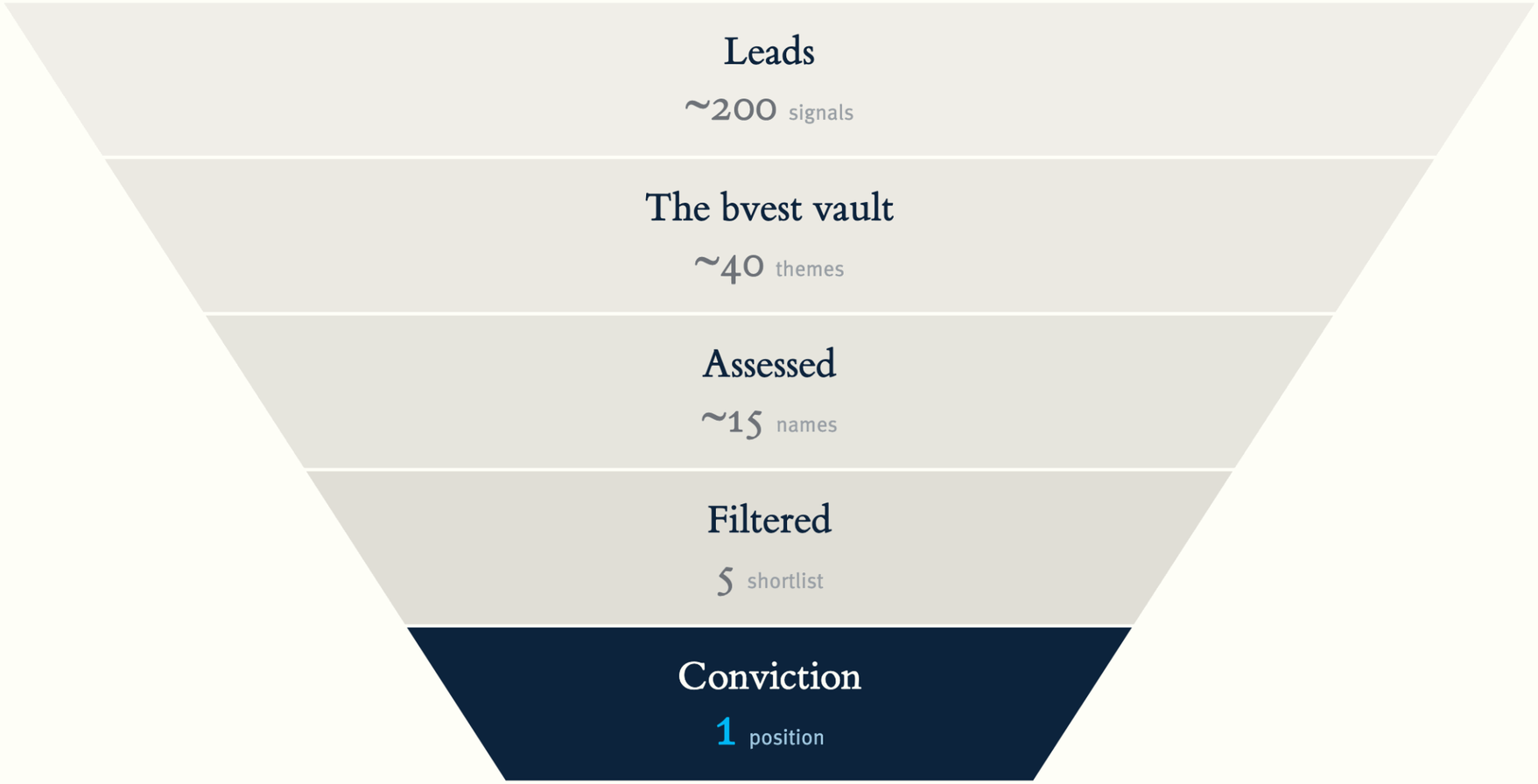


The investment decision

The same machine pointed at markets. A signal becomes a name; the name runs a gauntlet of filters and a deep dive; a committee takes the call.

Hundreds in, one position out

Every position starts as one of hundreds of signals. Each stage cuts hard, and only what clears them all earns capital.



Indicative attrition — the bvest vault and the book carry the live figures.

The bvest desk: seven seats and two producers

bvest is built as a hedge-fund org chart, one seat per skill. Work flows top-down — the macro call sets the ceiling, the funnel narrows the field, the committee votes and the PM sizes a single position — with two producers underneath.

MS

Macro strategist

REGIME & POSTURE

Breadth and macro signals into a conviction score and the exposure ceiling.

bvest-macro-strategist

ST

Strategist

THE IDEA FUNNEL

Podcasts, newsletters and scouts into themes and a ranked shortlist.

bvest-strategist

AN

Analyst

QUICK SCREEN

A stage read, a star check, Five Forces and a one-page tear sheet.

bvest-analyst

SA

Senior analyst

DEEP DIVE

The fiscal.ai databook into a 20-metric value-creation package.

bvest-senior-analyst

IC

Committee

THE VOTE

Five investor personas vote a verdict and a size.

bvest-ic

PM

Portfolio manager

THE BOOK

Hold, trim, add or close, with sizing and order specs.

bvest-pm

Q

Quant

PLACEHOLDER

A scaffolded seat, not yet active.

bvest-quant

RE

Render

PRODUCER

Tear-sheet markdown into a publication-quality page.

bvest-render

DB

Databook

PRODUCER

A per-ticker xlsx model that builds across deep dives.

bvest-databook

It starts with a signal

A Hermes scout watches X, podcasts and newsletters, scores what matters, and files it in the vault as a structured idea record — thesis, CANSLIM, evidence, a kill trigger.

Two or three names surface most days — already structured and scored, ready for the analyst rather than a pile of links to read.

X · podcasts · newsletters ⇢ the bvest vault

HERMES · INVESTMENT IDEAS · 2 NEW

TSEM

7.2 / 10

Tower Semiconductor

Specialty foundry moving from analog into **AI silicon-photonics**, backed by **$1.3bn** contracted 2027 SiPho revenue. Strong, but extended.

CANSLIM	C8 · A7 · N9 · S5 · L8 · I6 · M5
TECHNICAL	Stage 2, +18.8% on the 200-day, RS vs SPY +94.5% — extension risk high
EVIDENCE	Q1 revenue **$414m** +15%, gross profit +52%, Q2 guide $455m
KILL	SiPho contracted revenue delayed or amended down

A real scout digest, recreated. Mercado Libre arrived the same way — one of these that went all the way to the committee.

Macro sets the ceiling everything runs under

Before any single name, the macro strategist reads market breadth and the regime into one conviction score — and that score sets the exposure ceiling and the growth-versus-value tilt.

A broadening regime today — 70–90% net long. The Druckenmiller persona scores eight signals into the number.

The strategist decides where to focus

Signals become themes, themes become candidate niches. The Koch star rule keeps only the niches worth fishing in, and hands a ranked shortlist to the analyst.

LEADERSHIP

Number one

the leader in its niche by sales.

GROWTH

10%+ a year

a niche compounding for five years or more.

QUALITY

Profitable

a real moat, structurally able to stay that way — three ways distinct.

Clear all three and a theme becomes a **Star candidate** — a ranked shortlist for the analyst.

The analyst’s quick screen, on a real name

Each shortlisted name gets a fast read — the Koch star for quality, Perruna's stage analysis for timing, Five Forces for the moat — written up as a one-page tear sheet. Here is Mercado Libre.

KOCH · THE STAR RULE

A clear star

#1 in LatAm e-commerce, **4.5×** the runner-up; the niche has compounded ~16% a year for a decade and is profitable.

clears the quality gate

PERRUNA · STAGE ANALYSIS

Stage 4 — wait

~23% below the 200-day under a falling 40-week line, RS 31/100, off the low but well under the 2025 peak.

timing flagged for the committee

A clear star earns the deep dive — the timing question carried forward to the vote.

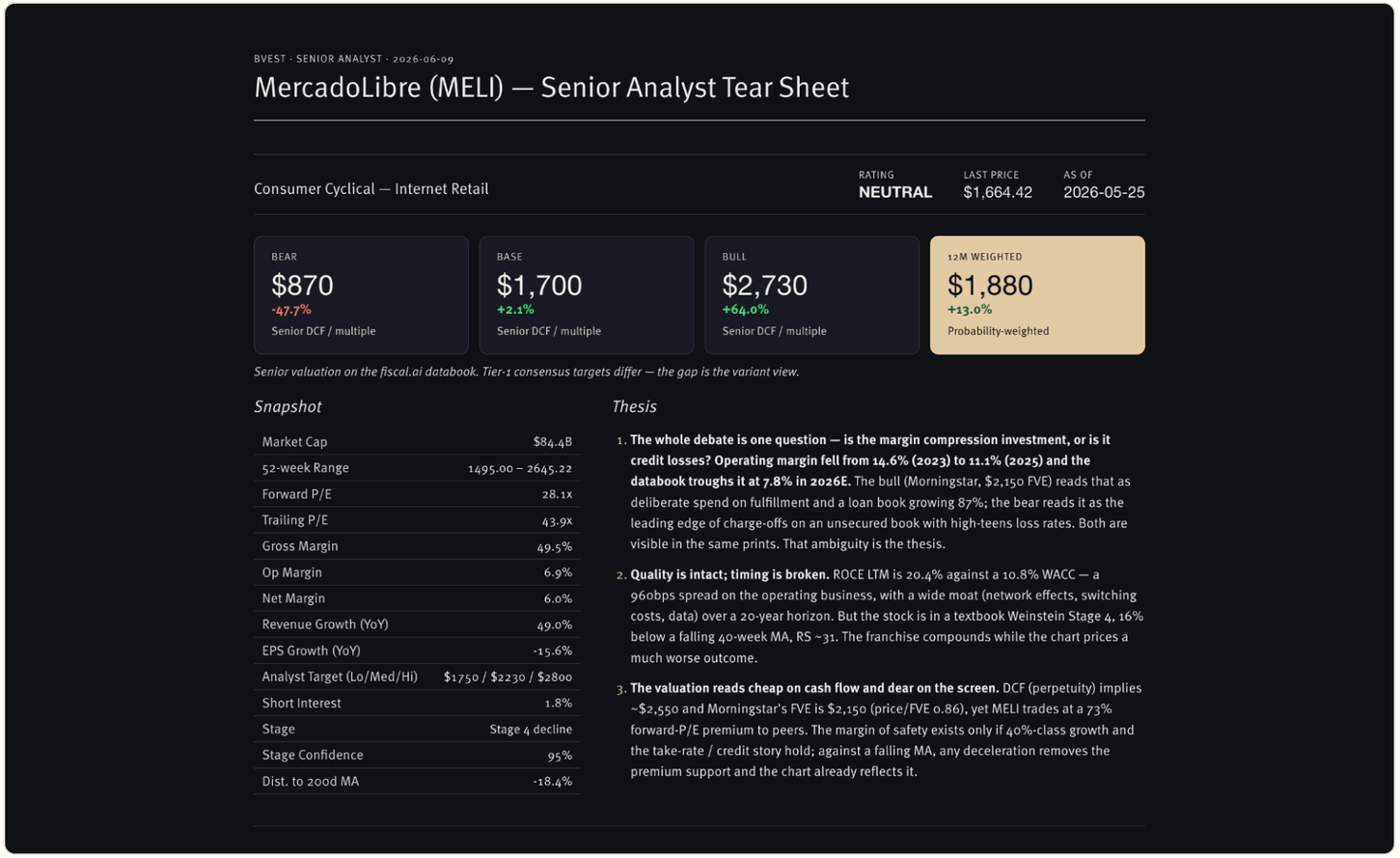
The senior analyst pulls the databook

A name that earns the deep dive goes to the senior analyst, who logs into fiscal.ai, pulls the full databook, and builds a 20-metric value-creation package with a technical read — the evidence the committee votes on.

ROIC against the cost of capital, the advantage period, a scenario tree — Mauboussin's value mechanics, run on real numbers.

The package ships as a rendered paper

The deep dive lands as a dark one-page tear sheet — bear, base and bull values, the probability-weighted target, the snapshot rail, the thesis — built by the render producer in the bvest house style.



BVEST · SENIOR ANALYST · 2026-06-09

MercadoLibre (MELI) — Senior Analyst Tear Sheet

Consumer Cyclical — Internet Retail

RATING	LAST PRICE	AS OF
NEUTRAL	$1,664.42	2026-05-25

BEAR	BASE	BULL	12M WEIGHTED
$870	$1,700	$2,730	$1,880
-47.7%	+2.1%	+64.0%	+13.0%
Senior DCF / multiple	Senior DCF / multiple	Senior DCF / multiple	Probability-weighted

Senior valuation on the fiscal.ai databook. Tier-1 consensus targets differ — the gap is the variant view.

Snapshot

Market Cap	$84.4B
52-week Range	1495.00 – 2645.22
Forward P/E	28.1x
Trailing P/E	43.9x
Gross Margin	49.5%
Op Margin	6.9%
Net Margin	6.0%
Revenue Growth (YoY)	49.0%
EPS Growth (YoY)	-15.6%
Analyst Target (Lo/Med/Hi)	$1750 / $2230 / $2800
Short Interest	1.8%
Stage	Stage 4 decline
Stage Confidence	95%
Dist. to 200d MA	-18.4%

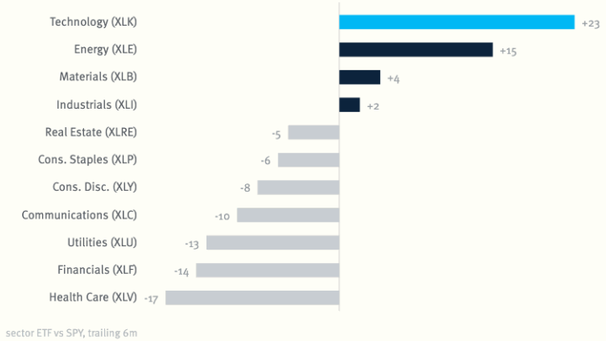
Thesis

1. **The whole debate is one question — is the margin compression investment, or is it credit losses? Operating margin fell from 14.6% (2023) to 11.1% (2025) and the databook troughs it at 7.8% in 2026E.** The bull (Morningstar, $2,150 FVE) reads that as deliberate spend on fulfillment and a loan book growing 87%; the bear reads it as the leading edge of charge-offs on an unsecured book with high-teens loss rates. Both are visible in the same prints. That ambiguity is the thesis.
2. **Quality is intact; timing is broken.** ROCE LTM is 20.4% against a 10.8% WACC — a 960bps spread on the operating business, with a wide moat (network effects, switching costs, data) over a 20-year horizon. But the stock is in a textbook Weinstein Stage 4, 16% below a falling 40-week MA, RS ~31. The franchise compounds while the chart prices a much worse outcome.
3. **The valuation reads cheap on cash flow and dear on the screen.** DCF (perpetuity) implies ~$2,550 and Morningstar's FVE is $2,150 (price/FVE 0.86), yet MELI trades at a 73% forward-P/E premium to peers. The margin of safety exists only if 40%-class growth and the take-rate / credit story hold; against a falling MA, any deceleration removes the premium support and the chart already reflects it.

The real artefact — equities/deep-dives/MELI-senior-paper.html, rendered 2026-06-09, screenshotted live from the vault.

Each committee member writes its own read

Before the vote, every persona reads the deep-dive package through its own lens — the macro regime, the star principle, the value mechanics, the tape — and writes its own paper. Here are the four reads on Mercado Libre, redrawn in the deck palette.

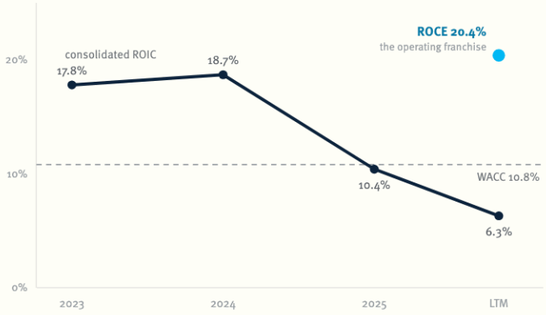


Macro regime · Druckenmiller

sector leadership — cyclicals lead, a risk-on read

Star principle · Koch

4.5× the runner-up · niche ~16%/yr · a clear star



Value mechanics · Mauboussin

ROCE 20.4% against a 10.8% cost of capital

The tape · Perruna

Stage 4 · ~23% below the 200-day

Redrawn from the live committee renders, 2026-06-09 — MELI-koch, MELI-mauboussin, MELI-perruna and the posture sector read.

Five lenses vote, the PM sizes the survivor

The committee does not re-run the numbers. It reads the analyst's package and argues. Each seat is modelled on a real investor's published thinking, drawn from that investor's own wiki.

Mauboussin
VALUE MECHANICS

ROIC against the cost of capital, the advantage period, expectations in the price.

Munger
JUDGEMENT

A latticework of models, inversion, the bias checklist, a default of wait.

Druckenmiller
MACRO

Regime, liquidity and the cycle. Sets the conviction and the ceiling.

Koch
STAR PRINCIPLE

Niche leadership times niche growth, and an 80/20 cut through the rest.

Perruna
TIMING

Stage analysis and a quality overlay. The enter, wait or exit read.

NEUTRAL · WAIT

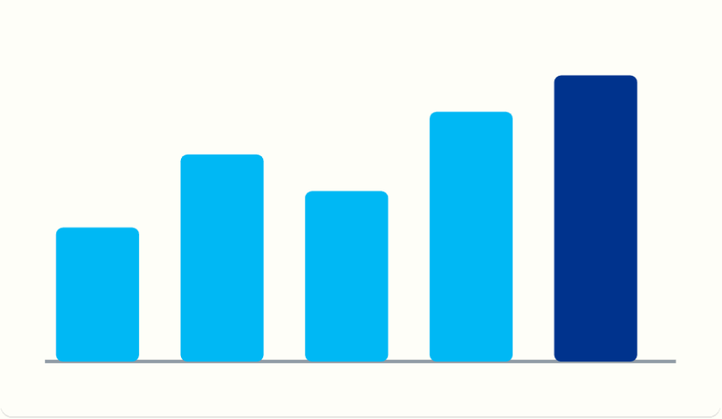
Mercado Libre. A clear #1 in LatAm e-commerce, and the perpetuity DCF sits ~50% above the price — yet the tape is Stage 4, ~23% below the 200-day, with the Brazil macro drag under review. Wait for a Stage-1 base to form.

Execute

*To a standard. **blook** turns data and structure into on-brand charts, decks and databooks, and **bvoice** turns notes into writing in one voice.*

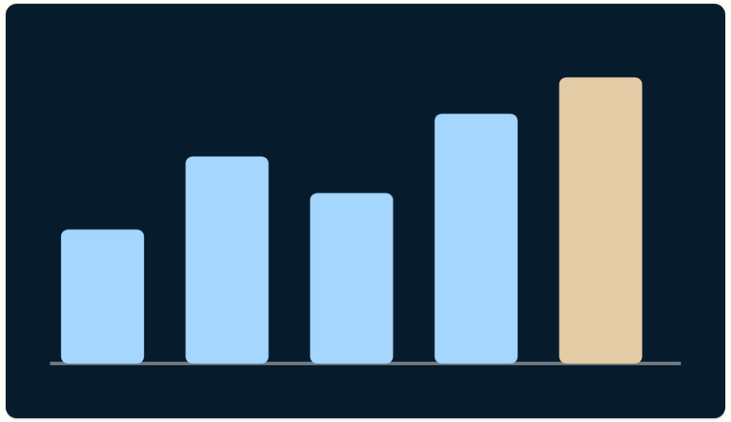
blook — one spec, any brand, any surface

The same chart, retinted to each brand and theme from one palette definition — then rendered to whatever surface the work needs.



KPMG

barnabyrobson.org — light

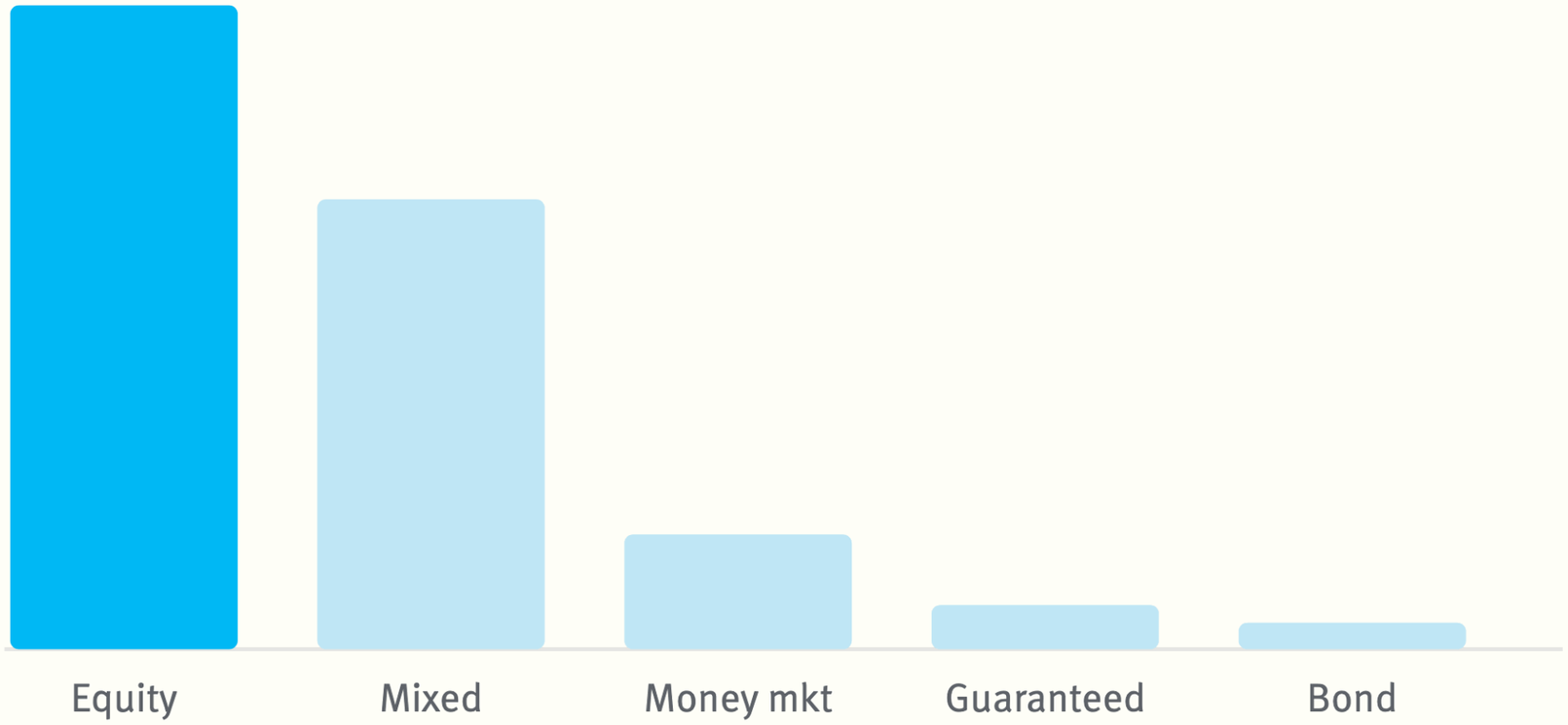


barnabyrobson.org — dark

RENDERS TO HTML · PDF · PowerPoint · Word · Website · Figma

HTML renders live — the reader can touch it

Of every surface blook renders to, HTML is the one that stays interactive. Hover or tap a class — the 447 MPF funds from the data-story, live in the deck.



☞ **Hover or tap a bar**

Equity

7.3%

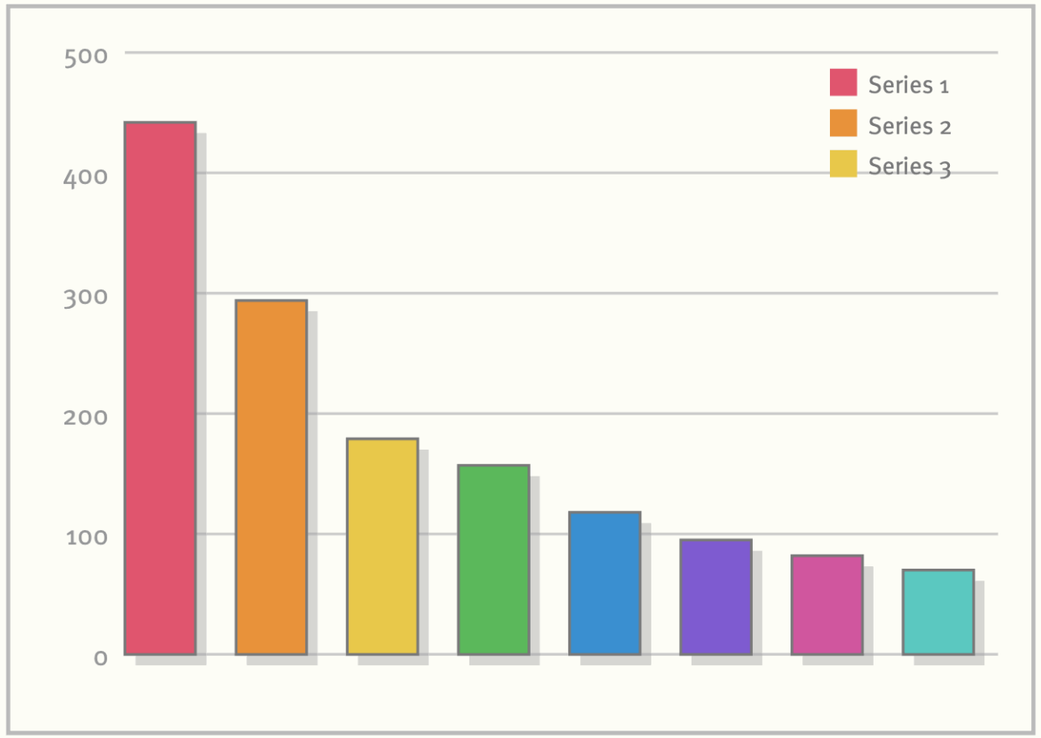
MEDIAN 10-YR RETURN · 145 FUNDS

blook-tufte — strip the chart to its data

A chart from the published MPF data-story, before and after the data-ink cut. Gridlines, shadows and rainbow go; the title becomes the finding and the accent marks only the schemes it names.

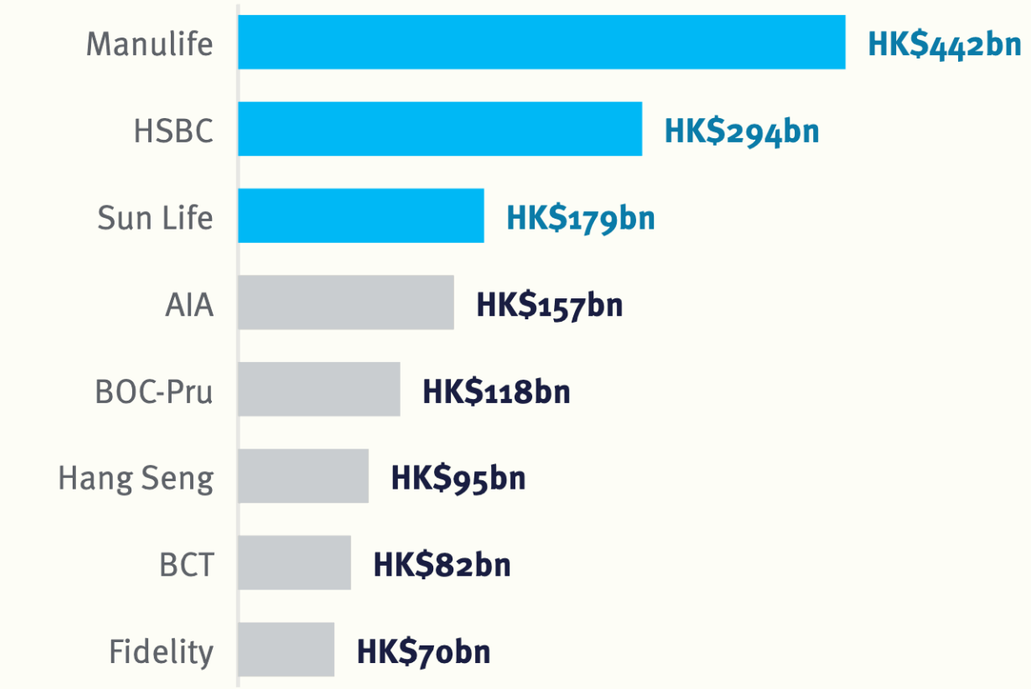
CHART-JUNK

Figure 3: AUM by Scheme (HK$bn)



BLOOK-TUFTE

Three schemes hold over half the pot



MPFA, schemes by asset size, 30 Apr 2026

One palette, routed everywhere it renders

Follow one hex. The gold premium signal is defined once, in the Blook Studio Figma file; **blook-palette** syncs it to tokens.yaml and regenerates every downstream mirror from that one definition.

tokens.yaml	`B-03:  hex: '#E3CBA6'  role: Brand Alt — gold premium signal`	**the canon**
wordpress-theme-partial.json	`"slug": "brand-gold", "color": "#E3CBA6"`	**the website**
bvest-render / render.py	`"B-03": "#E3CBA6",  # generated dark-canon block`	**tear sheets**
bvest-dark-blook.css	`--bv-brand-alt: #E3CBA6;  /* warm gold, selected signal */`	**Obsidian dark**

Hand edits to a generated slot are overwritten on the next sync — the hex lives only in the one definition. The gold accent bar in the retint three slides back is this token.

A deck-spec becomes a branded deck — this one

The same structured model behind a chart drives a whole presentation — one spec, rendered to on-brand slides in HTML and PDF. You are looking at one of them.

```
slides.push({
  sec: "Execute — blook",
  html: `
    <h2 class="head">blook-tufte
      — strip the chart
      to its data</h2>
    ${tufteClean()}
    ${foot("Execute")}
  `,
})
```

The spec two slides back, as written. blook-slides ships a render only on a green three-check gate: every spec number appears, every chart axis sums, nothing crosses the footer.

Notes to a published essay, in one voice

Writing runs as a pipeline with gates between the stages — each stage clears its gate before the next begins: storyboard, draft, review, rewrite, publish. The voice is inherited from a sample of my own published work.

1
Storyboard
A pile of notes and links into a signed-off claims list and arc.

2
First draft
Drafted in my voice from the storyboard and a voice anchor.

3
Review
The six-lens editorial gate, with before and after rewrites.

4
Rewrite
Findings applied section by section, in voice.

5
Publish
To WordPress, then cross-posted to X and LinkedIn.

This deck's copy ran the same pipeline.

The voice spec, enforced as you write

The spec is applied at write time, sentence by sentence; **bvoice-editorial-review** is the six-lens gate behind it, flagging the words and sentence-shapes that read as machine-written on every artefact a reader sees.

DRAFT — FLAGGED AT WRITE-TIME

We must ~~delve~~ into the ~~landscape~~ and ~~leverage~~ AI — agents arrive on machine cycles, not planning cycles — to ~~foster~~ a ~~robust~~, ~~comprehensive~~ strategy.

IN VOICE

AI now runs on machine cycles. Planning has to keep that pace.

Prose quality · AI-voice · Argument · Thought leadership · Structure · Visual data

COPE, end to end

One published data-story, start to finish — 447 funds scraped from the regulator, organised into a databook, charted by blook, and shipped as a live page on the site.

447 funds, scraped from the regulator and charted

Hong Kong's MPF regulator publishes every fund's returns and fees behind a slow web front-end. A Python script read the platform's own data endpoints, pulled all 447 funds into a databook, and the execute layer charted the result in HTML — published as a data-story on barnabyrobson.org.

01

Scrape

Python calls the MPFA platform's own data endpoints — no browser. 447 funds across 24 schemes, returns and fees.

CAPTURE

02

Databook

One row per fund, cross-checked and written to CSV and XLSX, the source documented on a read-me sheet.

ORGANISE

03

Chart

The figures drawn as in-page SVG, styled to the house palette by blook-tufte — dispersion strips, scatters, log bars.

EXECUTE

04

Publish

One self-contained HTML file, live at barnabyrobson.org/hong-kongs-mpf-visualised.

EXECUTE

THE UNIVERSE

447

funds across 24 MPF schemes, every one with returns and fees

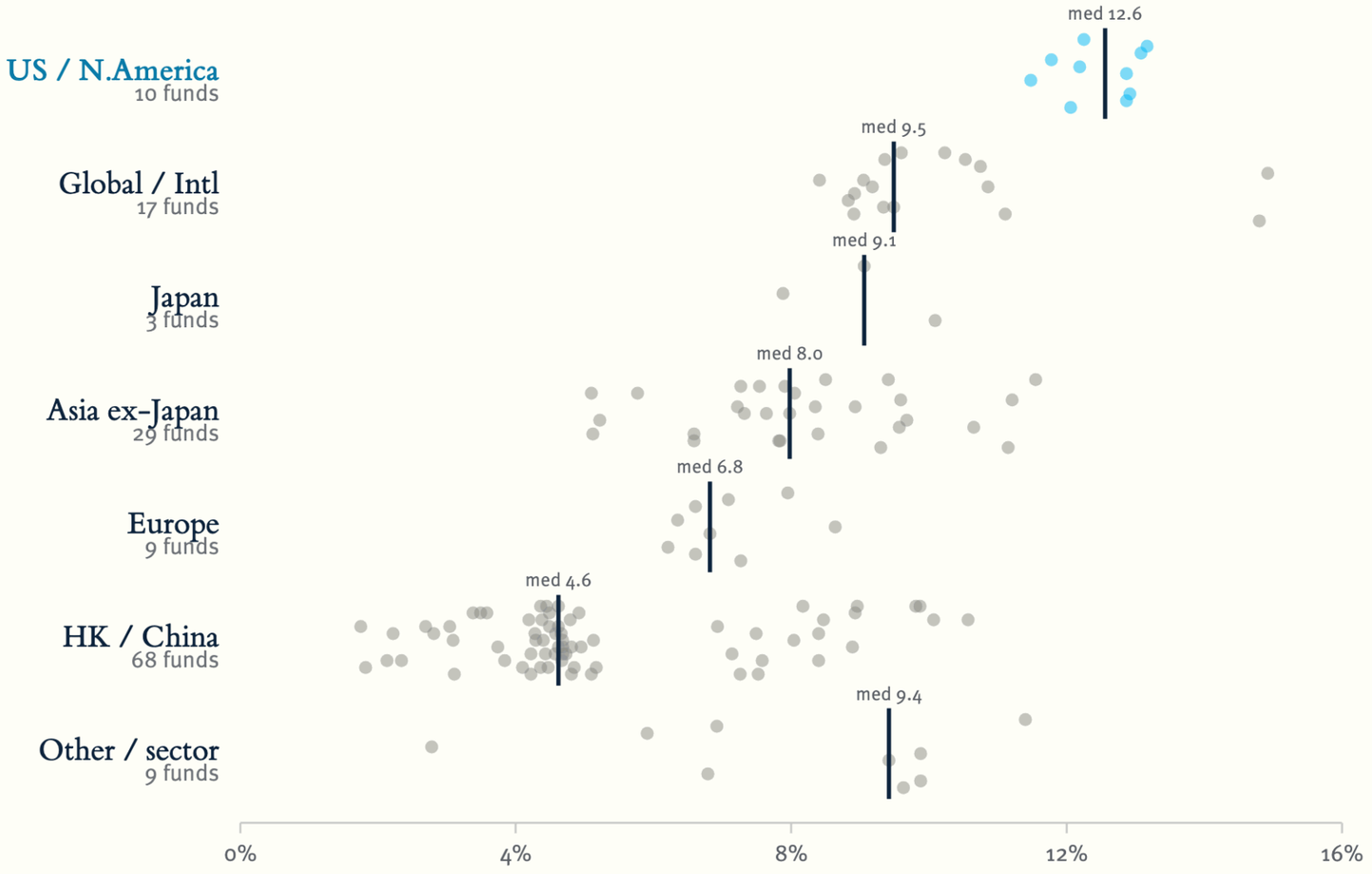
HK$1.63tn

the system mapped, as at 30 April 2026

Read straight from the regulator's platform, **no manual entry**.

Star-Crossed by Geography

Each dot is one equity fund's ten-year return, by region; the navy tick marks the median. US equity beat Hong Kong / China by **eight points a year for a decade** — and HK / China is the most-offered bucket on the menu.



From the 447-fund databook, drawn as SVG in HTML; published June 2026 at barnabyrobson.org/hong-kongs-mpf-visualised.

AI does not take your job. Someone running AI better than you does.

Runtime

The fleet that keeps itself alive — health, cron and repair across two servers, underneath all four phases.

Seven checks, every machine, on demand

brun-health runs a read-only pass across the whole estate — seven checks across four areas, each with an evidence code, appended to a running log so drift shows up over time.

MACS

The fleet

provisioning, the symlink farm, iCloud sync currency

SERVERS

Hetzner & Hermes

the gateways up, heartbeats fresh

SCHEDULE

Cron

every scheduled job's last success in window

KNOWLEDGE

Search & sessions

QMD indexes healthy, session logs current

Each run writes one row per check to the brun evidence log — the trail is the point.

Find the silent job, run the playbook

Two jobs keep the schedule honest. **brun-cron** inventories every timer across both servers and flags the ones that have gone quiet; **brun-repair** runs a per-incident playbook, gated so it refuses to act when the fault is not actually present.

BRUN-CRON · THE AUDIT

Every timer, one pass

- launchd, systemd, cron.d and Hermes jobs.json, read in one sweep.
- Each job's most recent success checked against a threshold.
- The silent ones flagged before they are missed.

BRUN-REPAIR · THE FIX

Playbook, gated

- A pre-check refuses to run when the fault is absent.
- Steps run; a post-check confirms the fix took.
- brun-github then publishes any skill changes, with a receipt.

Personas & superskills

Seven minds the stack calls as lenses, and the deep dives behind the committee vote.

Seven minds, each modelled from their own words

Cross-cutting personas the stack calls as lenses. Five sit on the investment committee; two stand alone for operating and people decisions. Each reads a corpus compiled from the real person's published work.

Mauboussin

VALUE MECHANICS

ROIC against the cost of capital, the advantage period, expectations in the price.

bvest-ic · senior analyst

Munger

LATTICEWORK JUDGEMENT

Models from three disciplines, inversion, the bias check, a default of wait.

bvest-ic · bdas

Druckenmiller

MACRO CONVICTION

Eight signals into a 0–100 score and the market pattern.

bvest macro

Koch

STAR PRINCIPLE

Niche leadership times niche growth, and an 80/20 cut.

bvest-ic · bdas

Perruna

STAGE TIMING

Stage analysis and a CANSLIM overlay: enter, wait or exit.

bvest-ic

Musk

EXECUTION COMPRESSION

Delete, simplify, accelerate, then automate — the Musk algorithm.

standalone

Horowitz

CULTURE & HARD CALLS

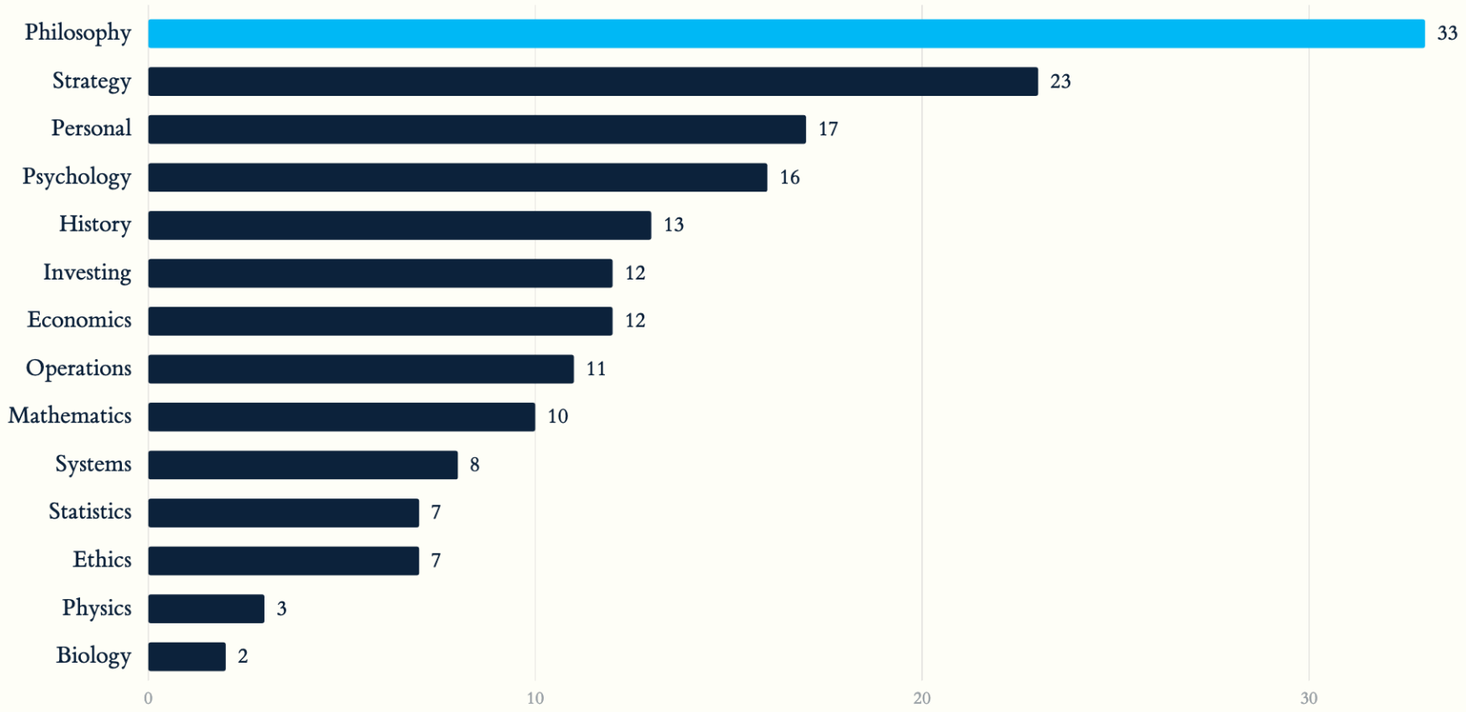
What a decision installs as culture; wartime versus peacetime.

standalone

183 mental models, across fourteen disciplines

MUNGER

The Munger seat reads a latticework compiled from Charlie Munger's own canon, *The Psychology of Human Misjudgement* at its centre. Before it weighs a thesis it pulls models from at least three disciplines, inverts the question, and runs the bias checklist.



THE LATTICEWORK

183

mental models, each tagged by discipline and by use-case

9

people pages — Buffett, Graham, Franklin, Darwin, Cialdini

The method: the 25 standard causes of misjudgement, inversion, lollapalooza convergence, and a demanded margin of safety.

The verdict: screaming yes, wait, or no — and the default is wait.

153 papers, compiled into one value-creation lens

The Mauboussin seat reads a corpus of his own writing — 153 source papers compiled into concept and people pages. It runs three modes: value creation, fundamental analysis, and forecasting.

MAUBOUSSIN

ROIC vs the cost of capital
value creation

Competitive advantage period
moat durability

Expectations investing
price as a forecast

Capital allocation
buybacks, shareholder yield

Base rates
the outside view

Skill versus luck
the continuum

Intangibles-adjusted ROIC
modern accounting

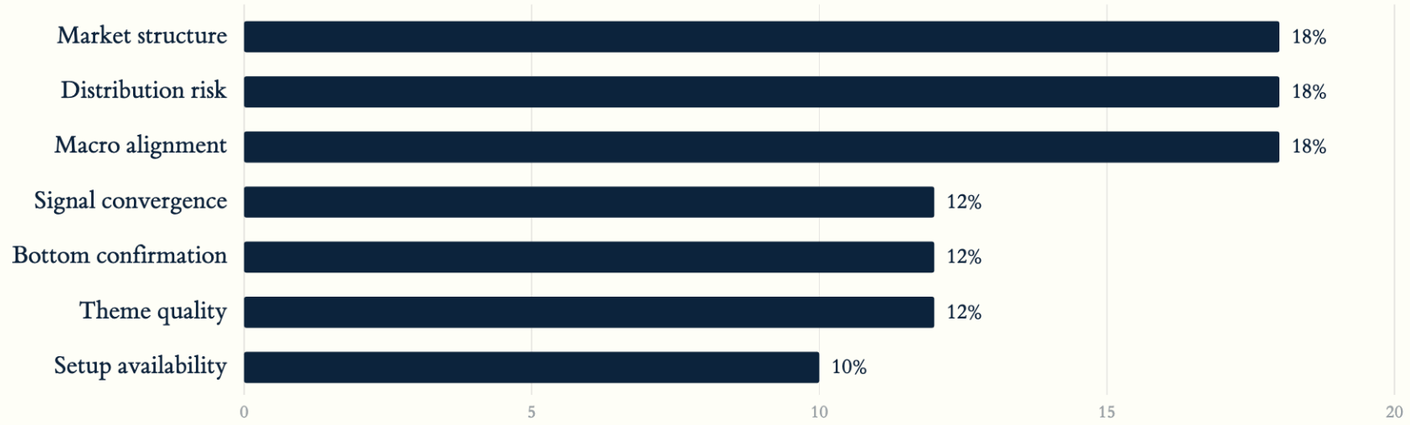
Market-implied expectations
the reverse DCF

160 concept pages and 58 people pages compiled from the 153 papers, retrieved live. The value-creation mode is the one the Mercado Libre deep dive ran.

Eight signals, seven weights, one conviction score

DRUCKENMILLER

The Druckenmiller seat reads up to eight upstream signal skills — breadth, sector uptrend, distribution risk, the macro regime, a bottom-confirmation trigger — and scores them across seven weighted components into a single 0–100 conviction number. That number sets the market pattern and the exposure ceiling everything else works under.



THE CONVICTION ENGINE

66

out of a hundred today — a broadening regime, 70–90% net long

4

market patterns — policy pivot, distortion, contrarian, wait

The rule: the score maps straight to an exposure band, from capital preservation at the bottom to a 90%-plus net-long fat pitch at the top.

Five sequential gates, then an 80/20 cut

KOCH

The Koch seat encodes two of Richard Koch's frameworks. The star principle runs a candidate through five gates in order — fall at any one and it drops out — while the 80/20 lens strips back whatever isn't carrying the result. Niche leadership times niche growth is the single best predictor it looks for.

01

The niche

A market you could lead: definable, profitable, and small enough to dominate.

DEFINABLE

02

Niche growth

Compounding above 10% a year, and durably across more than one cycle.

10%+ / YR

03

Leadership

Number one by sales, and by a wide margin over the runner-up.

#1 BY SALES

04

Isolating mechanisms

Network effects, switching costs, scale — the moats that hold the lead.

THE MOAT

05

Focus

Capital and attention on the core, with the 80/20 cut stripping the rest.

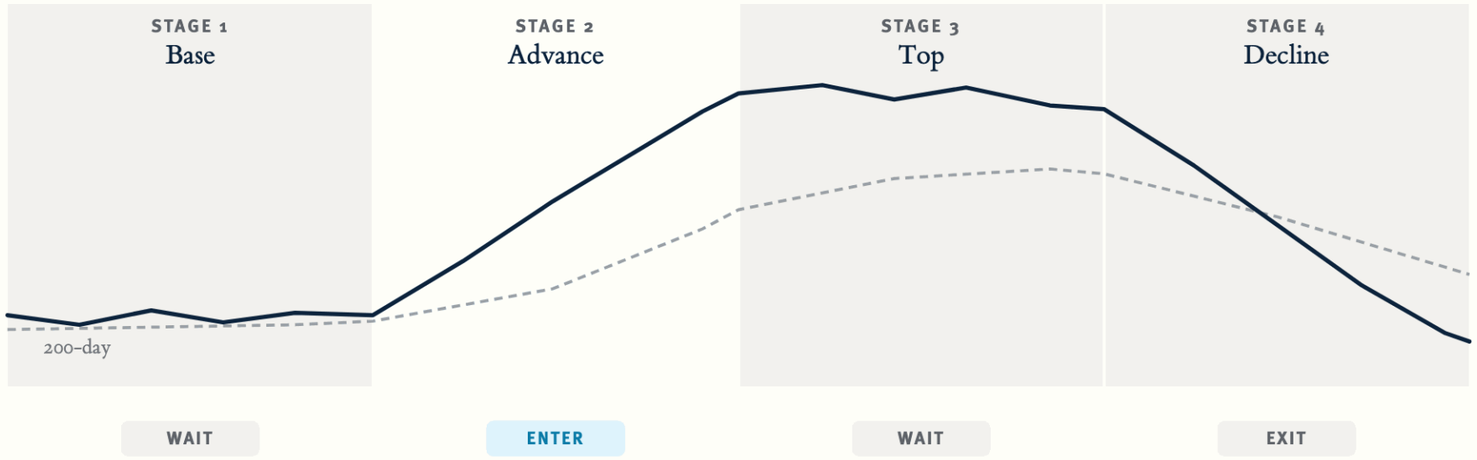
80/20 CUT

The lens that screened Mercado Libre through to a deep dive — a clear star, number one in its niche, the niche compounding around 16% a year.

Four stages, and a CANSLIM overlay

PERRUNA

The Perruna seat reads the chart through Stan Weinstein's stage analysis — base, advance, top, decline — then lays a CANSLIM quality score over it. It owns the timing the value seats leave open: enter, wait, or exit, with capital preservation ahead of any single trade.



THE OVERLAY

7

CANSLIM letters — current and annual earnings, leadership, supply, market direction

Stage 2

the only stage it will buy — a Stage 4 caps the position at zero

The gate: Stage 2, breadth above 55%, a CANSLIM score of nine or more, relative strength over 80 — all four, or it waits.

Five steps, in order — delete before you automate

MUSK

The Musk seat reads a corpus compiled from two biographies and the collected methods — 248 pages across 19 operating concepts. It runs one algorithm: a strict five-step sequence that fixes what a system is before it makes it faster. The order is the method.

01

Question requirements

Make them less dumb. Every requirement carries a named owner; a department cannot own one.

NAMED OWNER

02

Delete

Cut the part, the step, the report. If nothing ever gets added back, the cut was too timid.

CUT 10% TOO FAR

03

Simplify

Only what survives deletion. The schooling trap is polishing a thing that should be gone.

WHAT REMAINS

04

Accelerate

Speed the cycle once the shape is right. A faster wrong loop just makes more wrong output.

CYCLE TIME

05

Automate

Last of all. Tesla's error was robots built around a part that needed deleting.

EARNED LAST

The battery-mat case ran it backwards — automate, accelerate, optimise — then a test showed the mats were unnecessary; deleting them removed a production step and about two million dollars of robotics. The standalone seat the stack calls to compress a workflow before it automates one.

Two modes, opposite playbooks

HOROWITZ

The Horowitz seat reads a corpus from two books — *The Hard Thing About Hard Things* and *What You Do Is Who You Are* — 51 pages across eight concepts. It owns the people-and-culture seat: which operating mode the moment demands, and what each decision installs as culture.

	WARTIME	PEACETIME
PROTOCOL	Breaks protocol to win	Follows protocol to win
CULTURE	Lets the war define it	Defines it deliberately
CONFLICT	Heightens the contradictions	Builds broad buy-in
TOLERANCE	Completely intolerant	Tolerates deviation with effort
FOCUS	A speck of dust on the prime directive	Empowers people on the details

THE PEOPLE SEAT

51

wiki pages across two books, compiled into eight operating concepts

7

cultural-design techniques — from shocking rules to walking the talk

The method: culture is what you do — tolerate an off-culture act and you have set the new standard.

The seat: management, org design, and the manner of a hard decision — diagnose the mode before the tactic.

The AI toolkit

What I run underneath the stack — the models, editors and services the skills sit on.

What I run underneath the stack

The agents, the local compute, the knowledge base, the adopted skills with their authors, and the feeds.

TERMINAL & AGENTS

- **cmux** parallel agents
- **Claude Code** the stronger coder
- **Codex** shares one plan

LOCAL COMPUTE

- **Apple Silicon** 64GB+ runs models local
- **Ollama** small LLMs, offline
- **Playwright** browser automation
- **Marker** `datalab-to/marker`

KNOWLEDGE BASE

- **Obsidian** local-first markdown
- **Web Clipper** page ⇢ markdown
- **Dataview · Bases · Graph**

ADOPTED SKILLS

- **GitHub · ClawHub** find, fork, drop in
- **skill-manager · handoff** Artem Zhutov
- **gstack** `garrytan/gstack`
- **grill-me** `mattpocock/skills`
- **autoresearch** `karpathy/autoresearch`

DATA FEEDS

- **fiscal.ai** the deep-dive databooks
- **FRED** macro series
- **Yahoo Finance** quotes & history
- **FMP** 30yr fundamentals

The living version, with GitHub links and pricing: **barnabyrobson.org/resources/ai-tools**

The skills, family by family

All six families, skill by skill — bknow, bdas, bvest, blook, bvoice, brun.

bknow — the knowledge graph

Everything lands as markdown, files into a linked graph and stays coherent — with local search across every vault.

bknow-organise	Files the inbox, fixes frontmatter, repairs wikilinks, audits vault health against the Kepano rules.
bknow-process	Reweaves stale notes, cascades edits across backlinks, synthesises across the vault.
bknow-clippings	Turns web clippings into filed notes — categorised, tagged, backlinked, cleaned.
bknow-ingest	Converts PDFs, OneNote and email into clean vault markdown.
bknow-obsidian	Authors Bases, Excalidraw, Canvas and task-aware markdown the way Obsidian renders it.
bknow-qmd	Local search across 8,000+ notes — keyword, vector and rerank in one query.
bknow-sweep	The cron orchestrator: refresh indexes, process every inbox, validate links in one pass.
bknow-vault-setup	Scaffolds a new vault — folders, templates, CLAUDE.md, Base views.

8 skills · Organise — the vault the whole stack reads and writes.

bdas — the deal lifecycle

Frame the thesis, gather and classify the evidence, stress-test it with management, then synthesise to a client deliverable on the KPMG brand.

bdas-gtm	Opens the engagement: thesis, scope, storyboard, and the deal-context file every phase reads.
bdas-analyse	Structures the IRL, cross-foots VDR data, runs the analysis, classifies findings by deal impact.
bdas-fieldwork	Builds management Q&A from gaps, runs interviews, stress-tests the thesis, keeps a challenge register.
bdas-report	Compiles findings into the narrative — exec summary, client-lens review, voice scrub, debrief.
bdas-kpmg	Renders report, deck, note or databook on the KPMG brand contract.
bdas-cdd	Commercial DD: four to six parallel research agents, self-challenge, an industry read built on causal chains.
bdas-vdr	Drives real Chrome through a locked data room to markdown, with per-provider playbooks.
bdas-synergy-reviewer	Tests each synergy against ICAEW Tech 04/20 and a Big-Four evidence checklist.
bdas-prime	Reloads deal state after a gap and briefs the next move.
bdas-fieldwork/minutes	A raw transcript into a canonical KPMG meeting note, filed to the right vault.

10 skills, 4 KPMG formats · the Process phase of COPE.

bvest — the investment decision

Set the regime, turn signals into a shortlist, screen, deep-dive, vote at committee, then size and monitor the book — a hedge fund as a set of skills.

bvest-macro-strategist	Sets the regime: breadth and macro signals into a 0–100 conviction score and an exposure ceiling.
bvest-strategist	Front of the funnel: podcasts, newsletters and scouts into themes and a ranked shortlist.
bvest-analyst	Quick screen on a name — a stage read, a star check, Five Forces, a one-page tear sheet.
bvest-senior-analyst	Deep dive: the fiscal.ai databook into a 20-metric value-creation package.
bvest-ic	Convenes the five-persona committee to vote a verdict and a size.
bvest-pm	Runs the book: hold, trim, add or close, with sizing and order specs.
bvest-render	Tear-sheet markdown into a publication-quality page, via Figma or HTML.
bvest-databook	A per-ticker xlsx model that accumulates across deep dives.
bvest-maintain	Weekly sweep: refresh prices and stages, mark P&L, reconcile the watchlist.
bvest-trading-brief	Server-side cron: the day's market data into a short Telegram brief.

12 seats (10 active, 2 scaffolded), 25+ workflows · shown as an org chart in the investment-decision section.

blook — execute, on brand

Set the palette, capacity-map the template before a word is written, build and publish the visuals, drive Figma, then certify each render against a reference.

blook	Router, and the three contracts every output meets: brand, template, render.
blook-tufte	Tufte-style HTML and SVG charts and data stories, and the path that publishes them to the site.
blook-palette	Builds brand palettes from source and syncs the canonical tokens downstream.
blook-template-design	Capacity-maps each text slot before prose is written; syncs PowerPoint geometry to layout config.
blook-generate	One-off assets: portraits, text-in-image infographics, mind-maps, posters.
blook-qa	An autoloop that compares a render to a golden reference and re-renders until clean.
figma-use	The mandatory rules pass before any Figma Plugin API call.
figma-workflow	Patterns for iterating existing Figma files — page sweep, clone-and-retint.
figma-generate-design	Composes a Figma screen from an existing design system, section by section.
figma-generate-library	Builds a whole Figma design system from a codebase, in phases.

10 skills · the Execute phase of COPE. This deck and the tear sheets came out of blook.

bvoice — one voice, notes to published

Raw notes and audio become a storyboard, a voice-matched draft, a reviewed and rewritten piece, then a published, cross-posted essay.

bvoice-storyboard	Shapes a pile of notes, links and clips into a signed-off storyboard.
bvoice-first-draft	Drafts in my voice from the storyboard and a voice anchor.
bvoice-editorial-review	A six-lens review — prose, AI-voice, argument, structure, data — with before/after rewrites.
bvoice-rewrite	Applies the review section by section, any reorder gated on sign-off.
bvoice-crosspost	Fans an essay out to X and LinkedIn; browser-driven, never auto-publishes.
bvoice-wordpress-manager	Six WordPress jobs: publish, SEO, theme, links, indexing, performance.
bvoice-pre-publish-post-assistant	Suggests categories, tags and SEO from the live site taxonomy.
bvoice-cover-image	Builds the cover-image prompt and writes the path into frontmatter.
bvoice-doc-coauthoring	Co-authors specs and proposals to a stand-alone, reader-tested draft.
bvoice-podcast-transcript	Pulls a podcast transcript into a clean note with the quotes.

11 skills · the Execute phase of COPE. This deck's copy runs through bvoice, the voice spec enforced at write time.

brun — keep the fleet alive

Open with a briefing, check current state, audit the schedule or the registry, repair known faults, then publish any skill changes that came out of the session.

brun-prime	The day-one-back briefing: topology, provisioning drift, the week's incidents.
brun-health	A read-only seven-check pass across the Macs, servers, cron, search and sessions.
brun-cron	Inventories every scheduled job across launchd, systemd, cron.d and Hermes, and flags the silent ones.
brun-skills	Audits the Skills tree, rebuilds the symlink farm, checks the global rules for bloat.
brun-repair	Per-incident playbooks, each gated on a pre-check and confirmed by a post-check.
brun-github	Validates, pushes and promotes the Skills repo to the protected branch, with a receipt.

6 skills · Runtime, underneath all four phases.

Version 1 of this presentation was prepared on a Thursday afternoon, drawing on Barnaby's bstack skills, knowledge wikis and agentic session history. The document is written entirely in HTML.

Claude Code was the agent in this human–AI sandwich. Roughly **1m Opus 4.8 tokens** consumed — about **USD 75** over the API. Barnaby's time cost, at client rates: approx **USD 3,000**.

THE FRAMEWORK

barnabyrobson.org/on-making-things-up-changing-your-mind-and-cope

THE TOOLKIT

barnabyrobson.org/resources/ai-tools

THE PRECURSOR

“We are all going to have to become engineers”