

# Interview Agent

Where we are, and what happens next

From an office-hours session with Tong, 7 August 2026

Full record: DESIGN-saanpokkrong-2026-08-07.html

## 1. Where we are right now

### Today, Friday 7 August

The design is done and reviewed twice. Nothing has been built yet.  
Next thing to happen: one five-minute phone call, then Pete starts on Monday.  
Deadline for the build is Saturday 8 August, 23:59. The room is Thursday 13 August.

|    | Step               | What it produced                                                     | Status       |
|----|--------------------|----------------------------------------------------------------------|--------------|
| 1  | /office-hours      | Six decisions and this design                                        | Done today   |
| 2  | Adversarial review | Caught the wrong calendar, the CONTEXT.md leak, the dry-counter trap | Done today   |
| 3  | PBT cross-check    | Caught an inverted hub definition and three more routing errors      | Done today   |
| 4  | /plan-eng-review   | Assigns owners, locks the build                                      | <b>Next</b>  |
| 5  | /investigate       | Finds why the board takes 25 seconds                                 | After 4      |
| 6  | Build P0 to P3     | The working tool                                                     | Sat 23:59    |
| 7  | /qa                | Tests a silent participant and a confused one                        | Mon or Tue   |
| 8  | /review, /ship     | Landed and deployed                                                  | Before Wed   |
| 9  | The room           | ศาลปกครอง, 30 people                                                 | Thu 13 Aug   |
| 10 | /plan-ceo-review   | Settles the Q4 positioning question                                  | After 14 Aug |

Steps 1 to 3 are the shaded rows. Everything below them is unstated.

## 2. The short version

### Three things to carry out of this

The product is the room read. The conversation is how we collect it. That contradicts LD-027, which you locked two days ago, so LD-027 needs amending rather than quietly ignoring.  
The build is small and none of it is specific to this client. That is the only reason it earns two developer days for an 1,800 baht room.  
Nobody has ever paid for the read, and the room on 13 August cannot answer that, because a speaker booking has no budget line for a diagnosis.

### 3. What we decided

- D1** The product is an organisational diagnostic with a coaching-shaped front end. The unit is one room, not one person.
- D2** The wedge is the private answer paper cannot collect. A post-it wall in a court is a public act, so people write what is safe to be seen writing.
- D3** Three hours total. Our block runs between one and three. Other speakers fill the rest.
- D4** Reflection happens inside the workshop. People disperse instead of all speaking in one room.
- D5** Hard anonymity. Only the aggregate leaves. Said out loud at the top, engineered where it can be.
- D6** Approach C. The bot probes and types each person's card. The room clusters the cards and draws the links itself.

#### Your operating principle, which became the design

"PBT is this: we know the pathway, but for client they know better and we only provide good ingredient. And also menu of suggestion that really fits with their pathway. So we need to know what their network is, so that we can build a menu of fit interventions."

We supply the structure, the typing and the menu. The client supplies the content and picks. That is why Approach C is right and not merely convenient: hand a room a finished diagnosis and you have stopped being a supplier.

It also makes the network do real work. The order is network, then leverage point, then intervention family, then shortlist. So the wall has to be structured enough to route on.

### 4. The money, corrected

#### 600 baht an hour

Three hours at 600 is 1,800 baht. That is the government ค่าวิทยากร rate, not a consulting fee.

**The good half.** I first priced a failure as expensive, because REPORT.tex attaches 240,000 baht to this date. At 1,800 baht, an incoherent room read costs an awkward afternoon. Run the untested thing here.

**The serious half.** At speaker rates the buyer of the read is not in the room. A ค่าวิทยากร booking comes off a training budget, from whoever books speakers, with no mandate to act on a diagnosis of the court. So 13 August tests whether the read is coherent. It cannot test whether anyone will pay for it.

**The loose end.** The 240,000 baht figure is now unexplained. It is the only hard revenue number in this project's records and Lab planning may rest on it. Five minutes on the phone settles it.

#### What that does to the build

Nothing gets cut. The justification changes. Two developer days for an 1,800 baht room only works because every item survives past this client. The test for each one: does it still matter at UOB on 25 September, Siam Piwat from 15 September, and PoomJAI External in Q4? All eight do.

**Build nothing for this client. Build everything for September. Let 13 August be the rehearsal.**

### 5. Three tests, one run

**Does the tool collect?** UOB answered on 4 August, partly. Thirty-two of thirty-five people spoke and thirty got something back, so the mechanism holds in a live room on personal phones. The depth did

not hold. Only 5.7 per cent of answers carried a specific instance, the median answer was 35 characters, and 34 per cent of questions got no answer at all. Those defects are named and fixable.

**Is the room read coherent?** Untested. The one that exists we assembled by hand afterwards. Thirteen August is the first live attempt.

**Will anyone pay for the read?** No data. Every room so far has been sold as training, so the person paying is buying training. This is the test that decides whether there is a business, and nothing on the calendar is built to answer it.

## 6. What I think

**The asset is the corpus and the method, not the software.** The best thing to come out of 4 August was a structural read of an organisation, and what made it possible was thirty private accounts plus a body of OD knowledge to read them against. Any competent team can build a conversational agent. Few have four hundred practice atoms, a screened intervention library, and someone who can stand in a room and use them.

**The business risk is structural.** A tool that rides along free inside a training booking cannot discover its own price. Until someone writes a cheque for a diagnosis, this is an accessory to Ajarn's workshops. The fastest way to find out is not another build. Take the UOB structural finding back to UOB and ask for money for the follow-on. That conversation sits in MEETING-BRIEF.md marked undecided, and it is worth more than the September sprint.

**The positioning question answers itself if nobody answers it.** D1 chose the org instrument. LD-027 commits Q4 to a public launch with revenue secondary. Both cannot be the core product, and Q4 planning is happening now.

**We keep finding constraints in the room instead of before it.** The dominant UOB defect was visible in the code beforehand and shipped anyway, because every test had a participant who answered. The guard is already written in this project's own words: turn cost times step count against slot length, from telemetry we already had. Making that a habit is worth more than any single fix on the Saturday list.

## 7. The PBT check, and the error it caught

Tong asked for the design to be checked against Ong and Lee. Their 23 atoms were digested on 9 July, four days after this tool's PBT rigor pass, so none of them ever reached the code. This was their first contact with the design.

### The routing table was backwards

It read "a hub node (many inbound edges): high leverage, first target". The method says a hub influences many other nodes, which is out-degree, and that a node influenced by many others is over-supported and a poor first target.

So rows one and three named the same signature and gave opposite verdicts. The facilitator card built from it would have sent the room at the outcome and called it the cause.

server.py had it right all along: "hubs (out-degree), over-supported (in-degree)". The code was correct. The prose a human would work from is what drifted.

Three more errors in the same table. Moderators were being routed to an intervention family, and the method says they are not targets at all. In a court the moderators are statute, ระเบียบ, สตง. and อาวุโส, so a room writing interventions against those has produced a session that cannot move. The over-supported rule had lost the clause that makes it useful, "treat its upstream feeders first". And routing went from graph shape straight to an intervention family, skipping function entirely, which is the form-over-function thinking the method exists to reject. Routing is three hops: structure says which node, function says which intervention.

### The one that would have shown in the room

The interview asks two assessment questions. The card asserts four functional-analysis fields. Three of them have no question behind them, so the model fills the gap.

Anonymity makes that worse. An officer gives a thin answer, the card prints “หลีกเลี่ยงการถูกตำหนิ” as their short-term payoff, and they read it on a wall in their own court with no way to correct it without identifying themselves. No success criterion catches it. S7 counts clusters accepted without rewriting, so it would score this as validation.

The fix is small and it is in the list below: one more rung asking what happened next, their own words printed under each field, and blank fields left blank. A cluster with no antecedents printed on it is the room’s first question, not an embarrassment.

### Two relabels

The card should say สมมติฐาน, not root cause. The method licenses the word diagnosis only after a second pass against measures. A thirty-minute interview produces a working hypothesis, and the intake guide says so directly. You cannot argue with a root cause. You can argue with a hypothesis, and the argument is what this design says it wants.

The wall is not an idiographic network. Your own org-EEMM file settled this and the design was about to contradict it: “An org-EEMM is an evidence-weighted theory network, not a statistically-fitted idiographic network. An organization is not one individual. Nodes are real. Edges are weighted hypotheses, not computed correlations. Name this out loud in any deliverable.” The cards are evidence. The arrows the room draws are hypotheses, in a different colour.

### What survived

The person’s own words as the node label. The room drawing its own map, which the atoms treat as a psychological-safety move rather than a compromise. Physical cards, which the method considers a good medium. The four-rung ladder, which beats the method’s own five-parameter list for a room. And saying “screened hypothesis menu” instead of “proven interventions”, which is stricter than the method asks for.

## 8. The action layer

I told you the intervention half was already built. You rejected it: “this is not specific enough it’s ลอย and not actionable for each ppl na.” You were right. I judged the corpus at the consultant’s altitude, not the participant’s.

The four-family cheatsheet stops at labels like “Meeting Effectiveness”. A practice atom gets one rung lower. Neither is a thing a court officer does on Monday. So the cheatsheet routes, and the action layer has to be written.

Four rungs. The fourth is the one most decks leave out and the one a hierarchy lands on most. Worked example, cluster งานค้างเพราะรออนุมัติหลายชั้น:

|                            |                                                                                                                                                                                                                              |
|----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| คนเดียว ทำได้จันทร์<br>นี้ | บอกวันที่ตอนส่งเรื่อง. ทุกครั้งที่ส่งเรื่องให้หัวหน้า เขียนต่อท้ายบรรทัดเดียวว่า<br>ต้องการคำตอบภายในวันไหน และถ้าเลยวันนั้นจะเดินต่ออย่างไร<br>รู้ว่าได้ผลเมื่อ ภายใน 2 สัปดาห์ มีอย่างน้อย 1 เรื่องที่ได้คำตอบเร็วกว่าเดิม |
| ทำกับคู่                   | จับคู่เรื่องค้าง 15 นาที ทุกวันศุกร์. เอาเรื่องที่ค้างที่สุดของแต่ละคนมาวาง คนละ 1<br>เรื่อง ถามกันว่าติดที่ขึ้นไหน ไม่ต้องแก้<br>รู้ว่าได้ผลเมื่อ ครั้งที่ 3 เริ่มเห็นว่าส่วนใหญ่ติดขึ้นเดียวกัน                            |
| ทำกับทีม                   | ทำแผนที่ 1 เรื่อง ว่าผ่านมือใครบ้าง. เลือกเรื่องจริงที่เพิ่งเสร็จ เขียนบนกระดานว่า<br>ผ่านก็มีมือ แต่ละมือใช้เวลาเท่าไร ระบุตำแหน่ง ไม่ระบุชื่อ<br>รู้ว่าได้ผลเมื่อ ทีมชี้ได้ว่ามีขึ้นไหนที่ไม่เพิ่มอะไรเลย                  |
| ต้องผู้บริหารตัดสินใจ      | ยกวางเงินที่ไม่ต้องเสนอขึ้นไป. เสนอตัวเลขเดียวว่าเรื่องแบบไหนจบได้ที่ระดับไหน<br>พร้อมวิธีบันทึกให้ตรวจสอบได้<br>รู้ว่าได้ผลเมื่อ มีเรื่องประเภทหนึ่งที่ไม่ต้องขึ้นถึงชั้นบนอีก                                              |

Each card names the trigger, the visible move, the first sign within two weeks, the rung, and the source atom.

Write the generic deck, not a court deck. Six to eight root causes cover a bank, a retailer and a court alike: approval chains, workload concentration, cross-unit handoffs, information not reaching the top, senior-junior voice, backlog. That is about thirty cards, plus blanks for whatever the room surfaces that we did not predict.

## 9. Saturday 23:59, in order

Finishing Saturday leaves Monday and Tuesday for rehearsal and absorbs the 12 August holiday. Work top down and stop where the clock stops. Whatever falls off is desk work, never engineering.

|       | Item                                                                                                                                          | Who  | Est      |
|-------|-----------------------------------------------------------------------------------------------------------------------------------------------|------|----------|
| P0.1  | Phone call. Two days with a three-hour block, or one three-hour session? Is the court open Wednesday 12 August?                               | Tong | 5 min    |
| P0.2  | CONTEXT.md loads Lead+D's own internal problems as ORG BACK-GROUND at any client room. Set it per deployment, refuse to serve on the default. | Pete | 15 min   |
| P1.1  | Wait for the answer. No new card while the last question is unanswered and the mic is open.                                                   | Pete | 2 to 3 h |
| P1.2  | Clock gate. Warning at T minus 5, forced flush at T minus 0. Relabel the exit button, which reads หยุดพัก.                                    | Pete | 2 to 3 h |
| P1.3  | Never build a story before the first banked fragment.                                                                                         | Pete | 30 min   |
| P1.5a | Fifth ladder rung: ตอนที่มันติด คุณทำอย่างไรต่อครับ แล้วหลังจากนั้นเกิดอะไรขึ้น. Antecedent and consequence in one question.                  | Pete | 30 min   |
| P1.5b | Card prints their own words under each field, leaves unsupported fields blank, and says สมมติฐาน rather than root cause.                      | Pete | 1 h      |
| P1.5c | Card carries the malleable-versus-context call. Moderators print on second-colour stock.                                                      | Pete | 30 min   |
| P2.1  | Diagnose the 25.3 second board rebuild. Measure it.                                                                                           | Pete | 1 h      |

|      | Item                                                                                                                                                                       | Who  | Est      |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|----------|
| P2.2 | Cut the rebuild toward 8 seconds <b>and</b> make the dry counter time-based, same commit.                                                                                  | Pete | 2 to 4 h |
| P3.1 | Card generator, with redaction and a withheld path.                                                                                                                        | Pete | 3 to 4 h |
| P3.2 | A5 print stylesheet. This is the whole wall build.                                                                                                                         | Pete | 1 h      |
| P4   | Desk work in parallel: the bilingual family deck, the symptom-to-family page, the network rule card, new consent copy, the paper fallback, the twenty-minute candour call. | Tong | 4 to 5 h |

### Two things that are not negotiable

**P2.1 and P2.2 ship together or not at all.** The dry counter counts board rebuilds, not seconds. Cutting the rebuild from 25 seconds to 8 shrinks the window before the tool changes the subject from about 50 seconds to about 16. That pushes quiet people off their story three times faster, and quiet people are who the design needs talking.

**The wall is physical.** Print the cards on A5 and put them on a wall. A collaborative digital canvas does not exist today and is not a two-day build. Physical cards also make the fallback real, and a room that handles its own cards trusts the map more.

Out of scope before Saturday, said now rather than discovered at ten at night: the digital clustering wall, the challenge-board rebalance, the identity rewrite, the seventy-technique behaviour-change library, and September's instrumentation.

## 10. How we work with gstack

One loop, run in order. Each skill takes a document and hands you a better one.

| Skill              | Run it when                                         | What it does                                                                                                                            |
|--------------------|-----------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------|
| /office-hours      | An idea has no document yet                         | Forces the hard questions, produces the design. This is where today started.                                                            |
| /plan-eng-review   | Before anyone opens an editor                       | Locks architecture, tests, edge cases. It will also surface what this document cannot fix: nobody is assigned and there is no schedule. |
| /investigate       | Something is slow or broken and you do not know why | Root-cause debugging. This is P2.1, the 25.3 second rebuild.                                                                            |
| /spec              | A build item is still vague                         | Turns it into something Pete can implement without asking. Skip if P3.1 already reads clearly.                                          |
| /qa                | The build works on your machine                     | Tests it like a user. Here that means a participant who says nothing and one who echoes the card and says 33.                           |
| /review then /ship | Before it goes live                                 | Pre-landing pass, then deploy.                                                                                                          |
| /plan-ceo-review   | Scope or ambition is the open question              | Pressure-tests strategy. Right after 14 August, when the room read has one data point and Q4 needs a decision.                          |

## The rule of thumb

Before you build, run a plan review. When you are stuck, run /investigate. Before you ship, run /qa and /review. When the question is what to build rather than how, run /office-hours or /plan-ceo-review.

Today we did steps 1 to 3 of the ten in the first table. Step 4 is /plan-eng-review, and it should run before Pete starts on Monday.

## 11. Still open

The slot and the 240,000 baht line both close with one phone call. After that: whether the UOB structural finding goes back to UOB as paid follow-on work, whether LD-027 gets amended or quietly contradicted, whether September measures anything, what consent and data handling look like for a judicial client, who facilitates if this is the test of whether the tool works without Ajarn, and what one participant gets back for thirty minutes under an anonymous design.

That last one is small and easy to forget, and it is the one the thirty people in the room will notice.

Sources: data/uob-2026-08-04/ (manifest, audit findings, fixes, tool reflection, meeting brief), PASS-CRITERIA.md, server.py, wiki/concepts/pbt-network-approach.md, wiki/practices/pbt-network-building-8-steps.md, wiki/concepts/pbt-treatment-target-selection-criteria.md, wiki/practices/org-eemm-methodology.md, wiki/practices/od-intervention-family-cheatsheet.md, team/decisions.md. Every figure here was read out of those files or measured from the event logs.