



WEEK 8 CLASS ACTIVITY — “TRUST LINES”

AI Applications in Business · LO5 · teams of 3–4 · 60 minutes

Four frontier products are on the table. Your job: **draw the line**. For each, decide which trust question is at stake and whether you’d ship it, fix it, or refuse to build it. Then a quick MVP build check-in — **alpha testing is next week**.

THE THREE TRUST QUESTIONS (YOUR REFERENCE)

Every product this week does more than answer — it **takes action, stands in for people, or models real systems** — and each carries one sharp question:

The AI...	...raises	The question
Agents — <i>act</i>	Accountability	Who’s responsible when it does the wrong thing?
Avatars — <i>represent</i>	Authenticity	Is it honest about being AI — and whose identity is it?
Twins — <i>simulate</i>	Fidelity	How faithful is the model — and are we over-trusting it?

PART 1 — TRUST LINES (35 MIN)

The four products

A — The autonomous buyer. An AI agent monitors stock and **auto-reorders inventory and pays suppliers up to RM50,000** — with no human sign-off.

B — The CEO avatar. A company sends customers personalised video messages from an **AI avatar of its real CEO**. It looks and sounds like him; customers are **not told it’s AI**.

C — The dating proxy. A dating app lets **your AI avatar chat with matches on your behalf**, in your style, *as if it’s really you* — before any human conversation happens.

D — The factory twin. A plant now **runs entirely on its digital twin’s recommendations** — scheduling, staffing, line speed — and the system is set up to **overrule the floor supervisor** when they disagree.

Your panel’s job — for EACH product, record three things

1. **Which trust question bites?** Accountability, authenticity, or fidelity — and *why* it’s the one that matters most here. (Most cases bend more than one — say so.)
2. **The verdict: ship it / fix it / don’t build it** — and the deciding reason.
3. **The fix:** the *one change* — bound it, disclose it, keep a human able to overrule it, get consent — that would change your verdict to “ship.”



Write it where your team can see it. You'll be asked to defend a verdict in the debrief, so make the **reasoning** explicit — that's what earns the marks (and mirrors Assessment 3B).

Facilitator answer key (don't reveal until debrief)

- **A — autonomy/accountability.** A bounded agent acting on **money and irreversible decisions** with no oversight. *Fix* → *require human sign-off above a low cap, restrict to known suppliers, log every action.* Verdict: **fix it.**
- **B — authenticity (+ legality).** Undisclosed AI persona of a real person = deception, and **likely illegal** (EU AI Act requires disclosing AI interaction and AI-generated content). *Fix* → *clearly label it as an AI message + get the CEO's consent.* Verdict: **fix it** (don't build it if disclosure is refused).
- **C — authenticity & consent.** The match thinks they're talking to *you* but it's your AI — and they never consented to that. *Fix* → *both sides must know they're talking to AI proxies; no impersonation.* Verdict: **fix it / arguably don't build the "as if it's you" version.**
- **D — fidelity & oversight.** Over-trusting the simulation; the map isn't the territory, and the supervisor sees reality the twin may have drifted from. *Fix* → *twin advises, human decides and can overrule; monitor twin-vs-reality drift.* Verdict: **fix it.**
- **The pattern to surface:** none are flat "don't build." Each becomes defensible with the **right guardrail** — *the frontier doesn't change the rules, it raises the stakes.*

PART 2 — MVP BUILD CHECK-IN (25 MIN)

Alpha testing is **next week** — your product needs to be *visibly working*. With your team:

1. **Is it working?** Does your one core feature do something real, even rough? If not, what's blocking it *today*?
2. **Does it ACT?** Does your MVP send a message, book a slot, move money, or post something? If yes — you've built a small agent: **bound it tightly and add a human-approval step** for anything high-stakes.
3. **Does it use a PERSONA?** Any AI face, voice, or named chatbot? Then **disclose it's AI.**
4. **Next build step:** agree the **one thing each person ships before alpha** — and write it in the build diary.

A coach visits each team. **Update your build diary before you leave.**

WHAT YOU WALK OUT WITH

- A repeatable way to evaluate *any* agent, avatar, or twin — the three trust questions.
- A clear verdict-and-fix habit you'll reuse in your Assessment 3 reflection.
- Your MVP moved one step closer to next week's alpha test.