

Status: draft — grounds in `handoff/context/c4sight-schools-positioning-and-pricing.md`, `handoff/context/c4sight-schools-sessions-and-lessons.md` (Spine B, lower secondary), `deliverables/activities/interactive-activities-and-tool-spec.md` (A.7 Break the Bot, A.8 Spot the Fake), `deliverables/tangibles/curriculum-alignment-matrix.md` (§9, Years 7–8), and the Years 5–6 quality-bar lesson. Public preview pack — PDF-ready once the open questions at the end are settled.

Can You Trust It?

Deepfakes, confident lies, and what counts as evidence now — the AI conversations Years 7–9 actually need

Series 3 of 5 · Lower Secondary (Years 7–9) · Free preview — C4Sight

This pack is yours to keep and share with your staff. Inside: what the series is, one complete classroom challenge you can run this week with no cost and no special technology, and an honest account of what the full session adds.

What this series is

The core question: *Can you trust what you see, and can you trust what you make?*

Here is the honest bit first. If you want a free "how AI works" hour for Years 7–10, WADSIH (the WA Data Science Innovation Hub, at Curtin) runs a good one, and we're glad it exists. We don't compete with it.

This series is deliberately about everything that hour never touches — the part that actually keeps Year 9 coordinators up at night:

1. **What it actually is** — thirty seconds of honest demystifying. Not magic, not a brain: a pattern engine. Then we move on, because that's the part you can get for free.
2. **Where it breaks** — it makes things up, and it states false things *confidently*. Students watch a real example of a confident wrong answer and learn the word for it.
3. **Deepfakes and trust** — show, never make. A real-versus-fake challenge the room mostly cannot win. What does that do to "I saw it with my own eyes"?
4. **Your work** — using AI to help versus using it to cheat, where the actual line is, and what using it well looks like.
5. **Digital citizenship** — students will make and share this stuff. What is fair to make, what is not, and what the law is starting to say about deepfakes of real people.

The anchor activity of the full session is BUILD IT: the class writes an AI assistant's safety rules, then tries to break them. Some attempts fail. One works. The lesson lands physically: guardrails are just words someone chose, and the system underneath has no real understanding — which is why "the AI promised" is never a guarantee, and why trust has to come from somewhere else.

Our integrity stance, in one line: detectors do not work; design beats detection. We say it plainly to students and to staff, because it is true, and because a school that acts on a detector's verdict is

standing on ground that will not hold. Being the honest voice in the room is the whole point of this series.

Two people deliver it. A working AI practitioner who builds these systems for commercial clients, and a registered teacher who runs the room. Every other provider sends one person. The two-person model is not an overhead — it is the product.

The safety promise — in every C4Sight session, without exception

- No student devices.
- No student data entered into any system — ever.
- No student images captured, generated or uploaded.
- A school staff member present throughout.
- Where a camera is used, it frames objects the class holds up — never faces.
- Everything runs on our own device and connection, never your school network. The live classifier is fully offline.

Run this yourself: "Real or Fake?" — the trust challenge

A teacher-run discussion challenge for one class period. Everything below is enough to run it well from this pack alone. **The one rule that governs it: SHOW, never make.** You show pre-vetted examples from your screen. Students never generate a fake, never upload a photo, and nobody's image gets made into anything. The skill being taught is scepticism, not production.

Year levels	7–9
Time	25 minutes core (15-minute compressed version: rounds 0–2 plus a short debrief; full-lesson version: add the extended discussion)
Group	Whole class
Technology	Your device and the room screen only. No student devices needed. Round 0 needs no technology at all.
The point	Students feel their own confidence collapse — "it looks real" stops being proof, and "where did it come from?" becomes the reflex.

Before class: source your rounds

You curate a small set of real-versus-fake examples in advance. The table gives the example *types* and where to find each. You do not need all of them — Round 0 plus any two others makes a complete challenge.

Round	Example type	Where to source it	Notes
0 — the no-prep guarantee	A real quote vs a fake one, on paper	You write it yourself. Take a genuine quote from a text your class knows (novel, textbook, a reported news quote) and hand-write a plausible fake alternative. No AI, no technology involved.	This round always works, costs nothing, and rescues the lesson if any link fails.
1 — faces	AI-generated face vs real photograph	The University of Washington's public "Which Face Is Real?" project presents exactly this as a challenge — search <i>"Which Face Is Real University of Washington"</i> . AI-only faces can also be sourced from the well-known "This Person Does Not Exist" demonstration.	These are established public research/education resources built for this exact purpose.
2 — images	AI-generated image vs real news photograph	Several major news organisations have published "can you spot the AI image?" interactive quizzes. Search a trusted outlet's name (ABC, BBC, Guardian) plus <i>"AI image quiz"</i> .	Using a news organisation's own quiz means the vetting and the answers are done for you.
3 — voices	Synthetic voice vs real recording	News features on voice cloning often embed A/B audio samples of consenting subjects or public figures — search <i>"audio deepfake quiz"</i> plus a news outlet's name.	Audio reliably fools people <i>more</i> than images. Save it for late in the challenge.
4 — the confident lie	A documented AI "hallucination"	Reputable reporting on real incidents — the widely reported 2023 US court filing built on AI-invented case citations is the classic. Show the reporting, not a live tool.	This round shifts the lesson from "fake images" to "confident false text", which is where their homework meets the problem.
Reserve	One very hard image pair	Hold back your hardest pair in case the class is cruising.	Every strong class needs one humbling round.

Vet every example against this checklist before class:

- You know which item is which, and where both came from — provenance *is* the lesson, so you must be able to answer "how do we know?"
- No student, staff member, or local community member appears in anything.
- Real people appear only in already-public material: public figures, consenting subjects, published news photography.
- Nothing distressing, degrading or sexualised. Ever. If in doubt, swap it out.
- Check your links the week you run it — these resources move.

How to run it

1. **Set the frame and take the first count.** "One of the things I'm about to show you is real. One was made by a machine. You're going to vote." Then: *"Before we start — hands up if you reckon you could*

spot a fake." Count the hands and write the number on the tally sheet. Do not comment on it yet.

2. **Round 0 — the quote.** Read both quotes aloud (or display them). Class votes: which is the real one? Reveal, tally. This warms up the voting mechanic with zero technical risk.
3. **Run your sourced rounds, easiest first.** For each: show the pair (or the single item — "real or fake?"), vote by hands or by moving to sides of the room, record the votes on the tally sheet, reveal, record how many were fooled. Say nothing persuasive between rounds. Let the tally speak.
4. **Escalate.** Each round should be harder. By round three the room's confidence is visibly gone. Name it, once: *"An hour ago most of you had your hand up."*
5. **The audio twist.** If you sourced round 3, run it now. Ears fail more than eyes, and the room feels it.
6. **The confident lie.** Show the hallucination example. Make the link explicit: this is the same problem as the images — *confident* and *wrong* can live in the same sentence. A machine that invents a court case will invent a homework answer.
7. **Take the second count.** *"Hands up if you still reckon you could spot a fake."* Write it next to the first number. **The gap between the two numbers is the lesson.** Then run the debrief.

The tally sheet (print or copy to the board)

Before we start — "I could spot a fake": _____ hands

Round	What we showed	Voted REAL	Voted FAKE	The answer	How many fooled?
0 — the quote					
1					
2					
3					
4					

After the last round — "I could still spot a fake": _____ hands

The spoken teaching point (say it out loud, at the end)

"You just watched your own eyes and ears get it wrong, again and again. That's the point. 'I saw it' and 'I heard it' used to be proof. They're not anymore — so the skill now is asking where something came from, not whether it looks real."

Debrief questions

- Which round fooled the most of us? Why that one?
- "It looks real." After today, what does that actually prove?
- If seeing isn't believing anymore, what should be? (*Draw out: the source, the provenance, a second independent confirmation. "It was on the internet" is not a source.*)
- A photo of a classmate appears in a group chat and it's bad. What do you now know that changes what you do next? (*It may not be real; passing it on can be the harm; check before you spread.*)

- A voice note that sounds exactly like your mum asks you to do something odd. What's the move? (*Verify on another channel. The voice is no longer the proof.*)

If it stalls

- **The class aces the image rounds.** Good — switch to audio: *"Your eyes are trained. Your ears aren't."* Then play the reserve pair.
- **A link is dead or the screen fails.** Run Round 0 and go straight to the debrief. The zero-technology round carries the whole teaching point on its own.
- **A student goes nihilist — "so nothing's real, why care?"** Don't argue; reframe. Plenty is real, and provenance finds it: checking sources is how journalists, courts and scientists earn trust for a living. The skill is *checkable* trust, not blanket doubt.
- **Someone asks "can we make one?"** This is the moment the SHOW-never-make rule earns its keep: no — making a fake of a real person is a consent violation before it is anything else, the law around it is moving fast, and the skill that will serve them is spotting and checking, not producing. (The full session covers what the law is starting to say, at an age-appropriate level, kept current.)

What the full session adds (honestly)

This challenge is real value, and you can run it forever without us. Here is what changes when we're in the room — and it is a genuinely different lesson:

In this pack	In the full session
A challenge you source yourself	A pre-vetted, escalating deck we build and keep current — including the synthetic-audio rounds that are hardest to source well — engineered so the confidence collapse lands on schedule
Hallucination <i>described</i> , via reporting	Hallucination <i>demonstrated live</i> on the presenter's own machine — the room watches a confident wrong answer happen
The discussion of trust	The BUILD IT moment: your students write an AI assistant's guardrails on the big screen, attack them, and watch one bypass work. "Guardrails are just words" stops being a claim and becomes something they did
The integrity line, stated	The integrity conversation, had properly: why detectors do not work, what the actual line between help and cheating is, and what using AI well looks like for <i>their</i> assessments
The law, mentioned carefully	The deepfake-and-consent law beat delivered by a UWA JD (law) student — current, credible, age-appropriate
One teacher running one activity	Two presenters — a working AI practitioner and a registered teacher — running the full five-beat arc
This pack	A teacher resource pack you keep: pre-session primer, teacher-facing run sheet, framed video links, post-session activities and the curriculum mapping

Where it sits in the curriculum

An incursion **connects to and enriches** the curriculum — it does not deliver a descriptor to the achievement standard, and we won't pretend otherwise. For Years 7–8 the series speaks most directly to (Australian Curriculum v9, Digital Technologies):

Code	Descriptor	How this series connects
AC9TDI8K03	investigate how digital systems represent text, image and audio data using integers	Images and voices are just data — which is exactly why they can be synthesised, and exactly why "it looks real" proves nothing.
AC9TDI8P14	investigate and manage the digital footprint existing systems and student solutions collect, and assess if the data is essential to their purpose	The digital-citizenship beat: what you make, share and feed into systems, and what it costs you.
AC9TDI8P10	evaluate existing and student solutions against the design criteria, user stories and possible future impact	Judging what is fair to make and share, and its impact on real people — the evaluation is the ethics.

The heavier lifting sits in the **general capabilities** at this band (Level 5): **Digital Literacy** (locating and evaluating information — recognising credible versus unreliable sources is the deepfake skill), **Ethical Understanding**, and **Critical and Creative Thinking**. The series is also built to be consistent with the **Australian Framework for Generative AI in Schools** — particularly its *Human and Social Wellbeing*, *Transparency*, and *Privacy, Security and Safety* principles.

The full descriptor-by-descriptor mapping lives in the C4Sight Curriculum Alignment Matrix, which every student-facing booking receives.

Bring the full series to your school

All prices exclude GST · Perth metro · two presenters unless noted.

Format	What it is	Reaches	Price
Single Year Group	2 × 60 min sessions (one presenter)	Up to 60	\$950
Half Day	3 sessions	Up to 90	\$1,750
Full Day	5 sessions + a free 45-min after-school staff briefing	Up to 150	\$2,850
Term Program	6 weekly sessions with one year group, ending in a student showcase	One year group	\$7,500

Single Year Group, Half Day and Full Day bookings include the teacher resource pack. Many schools pair a student day with **Staff PD (\$2,200, or \$3,400 extended)** — it comes out of a different budget line, and the integrity conversation works best when staff and students hear the same honest story.

Or start free

Book the free 30-minute Taster. One class or a short assembly, one presenter, no cost. The rules: one per school, ever; a decision-maker (Principal, Deputy, HOLA or Digital Technologies coordinator) in the room; and it ends with a scheduled conversation about where AI fits at your school.

→ caleb@c4studios.com.au · c4studios.com.au

Who you'd be booking. We don't ask you to take claims on faith. What you can check for yourself: we run a Perth studio building AI automation for commercial clients — real, current work. Our practitioner is a UWA JD (law) student, so the ethics, consent and regulation content is genuine, not borrowed. We produce the C4Sight explainer series — public work you can watch before you book. And we deliver alongside a TRBWA-registered teacher.

C4Sight is a C4 Studios venture · Caleb Scott · ABN 26 477 523 297 · Perth, WA

Open questions for Caleb

1. **Do we host a starter pair?** The pack coaches teachers to source their own examples (with Round 0 as the no-prep guarantee). The stronger lead magnet would include one C4-vetted image pair hosted on the download page — but that deck carries reputational weight and ties into the activities doc's open question about who curates Spot the Fake material. Source-it-yourself (as written), or add one hosted C4 pair?
2. **Naming WADSIH in a public pack.** The capability one-pager already names WADSIH approvingly, and "we don't compete with the free hour — this is what it never touches" is the line a Deputy remembers. Confirm you're comfortable with it in a downloadable artifact, or I soften to "the free options".
3. **Years 7–8 code verification.** The three codes above are quoted from the alignment matrix (§9), which the editorial review flags as still needing the pre-print verification pass against the live ACARA v9 text. Standing task before this pack is exported to PDF.