DAY 36–45 · Part 4 · What to Watch Out For

Part 4 Review — What to Watch Out For

From protecting your information to AI law · with the real terms · DAY 36-45

No test — a review to read through

Once you type it in, it's out of your hands

1. Assume that whatever you type into AI leaves your hands the moment you send it.

2. Leave out the things that **can pinpoint you** — name, address, phone number — and the answer comes out the same.

3. Your friend's face or phone number isn't yours to post anywhere.

Pieces point to you when they're combined

1. "A 9th-grade boy" on its own could be hundreds of thousands of people.

2. Add a school and a homeroom and it's thirty; add tutoring times and it's **one person**.

3. Privacy laws work this way too: if it can easily be combined to tell who you are, it counts as personal.

Three things to check before you send

1. One: can it pinpoint me? Two: is it mine? Three: is it a password?

2. Passwords gain you nothing and risk everything, so they're **never OK, no matter what**.

3. Whether what you type is used for training, and how to turn that off, is written in the official help pages.

"Copying" got one more box to check

1. "Copying" used to be judged only by how much the result looked like the original.

2. Now there's a second question: **what did it learn from before it made anything**?

3. Countries answer it differently and court cases are still open, so learn the issue instead of a rule.

One more box to check, that's all

1. The old standard didn't disappear — a result that looks alike is still a problem.
2. What's new is a box called **what it learned from**.
3. That's why people now argue even when the results look nothing alike.

Lines drawn differently in each country

1. Europe let rights holders mark their work in advance to say "don't train on this."

2. Japan's government wrote down that even its "training is free" clause has limits.

3. The US Copyright Office published a report and still **didn't nail down a conclusion** — not knowing is the normal state here.

Your eyes can't tell — the source can

1. You can't tell by looking — the old trick of counting fingers barely works anymore.
2. Instead of picking apart the screen, follow **where it came from**.
3. Is it where it should be, and did anyone else film it too?

Clues that stopped working, clues that work now

1. The clues that stopped working all look inside the screen: fingers, blinking, image quality.

2. The clues that work now all look **outside the screen**: source, other angles, a sudden demand.

3. Inside the screen keeps getting more perfect, so move your eyes outside it.

Labels are starting to come from the makers

1. A way to plant invisible marks in text was published in a journal in 2024.

2. There's also an open standard for recording in a photo or video where and how it was made.

3. Europe has put rules for labeling AI-made content and disclosing deepfakes **into law**.

It isn't hate, it's a lean

1. AI doesn't decide to dislike anyone.

2. If the data it learned from leans one way, **its answers lean the same way** — that's all.

3. So when a question sorts or picks people, add one more: "What about other cases?"

What a 2018 study actually measured

1. Programs that guess a person's sex from a face photo were tested by skin tone and sex.

2. The error rate was highest for **darker-skinned women's faces**.

3. It was lowest for lighter-skinned men's faces.

A lean flows in from the front

1. If the data collected leans, training takes the bigger side as default, and answers go wrong for the rest.

2. A US standards agency also measured demographic differences in face recognition at scale and published a report.

3. Fixing AI is hard, but **asking once more at the end** is something you can do today.

One line of credit is a fence around your share

1. Saying you used AI isn't a confession — it's a name tag.
2. Stay quiet and **even the parts you did well** get doubted.
3. Split the assignment into pieces and write one line on who did each. That's it.

How to write that one line

1. Split the assignment into topic, sources, draft, polish and fact-check.
2. Mark each piece with whether you did it or AI did.
3. Which piece, how you used it, and what you changed — **those three** make the line complete.

Marks machines can read exist too

1. Google puts invisible watermarks into what its own models produce.

2. The European Union has written into law that AI-made content must carry **a mark machines can read**.

3. Exactly when that kicks in is still being adjusted.

Did anything stay in your head?

1. The test isn't a long rulebook, it's one sentence: did anything stay in your head?

2. Hand the whole thing over and **only the finished product is left** — you can't explain it.

3. Ask only about the stuck spots and rewrite it yourself, and you keep the work and the skill.

Four steps when homework lands

1. When you get homework, first ask what it wants you to learn.

2. Then try it with your own brain, even for 5 minutes.

3. If you're stuck, ask **for a hint, not the answer**, and write the last version in your own words.

Same tool, different place

1. Handing the whole thing over shows up fast when you present or sit a test — you can't explain it.

2. UNESCO also published official guidance on using AI in school.

3. In the same survey, **fewer than 10% of schools had formal guidance** — which is why you set your own line.

Spotting it is about order, not a sharp eye

1. Spotting fake news is about order, not about having a sharp eye.

2. Who said it first, when it happened, whether it's anywhere else — **just those three**, in turn.

3. If even one of them is empty, stop there. Stopping is a decision too.

Three boxes in under 1 minute

1. Box one is the source — trace it back to whoever said it first.

2. Box two is the date — old stories get passed around as if they happened today.

3. The strongest is box three: **did anyone else report it separately and say the same thing**?

Why "it looks convincing" stopped counting

1. AI can now produce text and photos in seconds.
2. So **it looks convincing** no longer works as a reason.
3. That's exactly why the habit is to trust clues rather than your eyes.

Hang up, then call back on a number you know

1. AI scams follow the same steps every time: copy, pretend, rush, demand.

2. Stopping it is almost boring — **hang up, then call back on a number you already knew**.

3. If it really is your family, calling back changes nothing.

Three things scams aim for, three that break them

1. Scams aim for a voice you know, the rush of "right now," and making you decide alone.
2. Flip each one: hang up and call back, pause a beat, tell an adult nearby.
3. Don't try to judge it by ear — **it's easier to change who places the call**.

A warning straight from investigators

1. US investigators issued an official warning that AI-made voices, photos and videos are used in scams.

2. A short voice clip is enough to copy someone closely.

3. The more urgent it feels, the faster a pause gets you there — **a real emergency is still urgent in 5 minutes**.

Are you the one deciding how much to hand over?

1. Nobody is saying don't use it — the whole point is whether you decide how much to hand over.

2. It's not 0 and not 100: **the most stays with you in the middle**.

3. Give it 5 minutes with your own brain first, then hand it over.

What a 2025 experiment measured

1. A research team ran a writing experiment and looked at the people who handed more to AI.

2. They reported that group **couldn't remember sentences they had just written** — even though the writing was fine.

3. It's a preliminary study still under review, so don't memorize it as the final word.

Order matters more than share

1. At 0% it's slow but it sticks; at 100% it's done but nothing sticks.

2. The middle bar is longest for one reason: **it leaves you the work and the skill**.

3. There are public self-check sites for how much you're leaning on your devices, too.

The bigger the risk, the stricter the rules

1. AI has laws too. But they don't treat every AI the same way.

2. They sort uses **into boxes by how risky they are** and set different rules for each.

3. South Korea's AI Basic Act took effect in January 2026.

Four boxes in the EU AI Act

1. Banned outright, high-risk, must disclose, minimal — four boxes.

2. The banned box can't be used at all; the high-risk box has to pass checks and be supervised by people.

3. The must-disclose box means **saying it's AI**, and the minimal box has no special rules.

The law looks at the use, not the technology

1. The same face recognition goes in a low box when it unlocks your phone.
2. Use it to pick people out secretly and it moves **to a top box**.
3. The rule that stands out in South Korea's AI Basic Act: content made by AI has to be labeled as made by AI.

It isn't memorizing rules,
it's keeping the checking box on your side.



Part 5 looks inside.