

Teaching AI Safety in Schools PD

1) Teacher Reality & Understanding Student AI Interactions

Where are students already interacting with AI? (list tools, apps, features, or situations)

One moment you've seen students trust a tool too quickly (what happened?)

Patterns we noticed (circle / note): privacy / oversharing / rumors / homework accuracy / scams / deepfakes / other: _____

2) Why AI Can Feel Like a Person to Students

Why students may assume "it's a friend" (star what you see most):

Warm/personal tone	Instant responses anytime	Mirrors emotions/language	Seems to remember preferences
Avatars/names/personalities		Encourages continuous conversation	

Examples we see at our school (specific moments, phrases, or behaviors):

Tool categories students may be using (check what applies):

- Character-style chat apps (role-play/relationships)
- Social media chat assistants
- General AI chatbots (homework/advice/entertainment)
- AI "companions" marketed around conversation
- Game-based bots/NPC chat and server bots
- Writing helpers/study assistants

Staff alignment message (in your own words):

When a student says, "It told me..." our consistent response should be:

Trusted humans/sources we redirect to (names/roles/systems):

3) Personal Identifying Information + The Billboard Test

Definition (write it in a student-friendly way):

What students overshare most often at our school (top 5):

What feels "normal" to students but shouldn't be online:

Billboard Test wording we will teach:

Corrections we want consistent across classrooms (what do we say/do?):

4) App Permissions + Digital Footprints

Permissions students click without thinking (examples):

How we'll support this when devices are used at home too:

Student-Friendly Definition of Digital Footprint:

5) Misinformation + Deepfakes

AI misinformation students actually see here (check and add examples):

- Fake "breaking news" posts/screenshots
- Deepfake video/voice
- AI images presented as real photos
- Confident-but-wrong homework help
- "Health/safety" claims
- Giveaways/free rewards scams
- Fake school rumors:
- Bullying/framing content:

Top 2 we see most:

1. _____
2. _____

Verification routine we will teach (write final wording): Pause, Source, Confirm (2 checks):

Schoolwide "before you share" expectation(s) (1-2):

Student script we want them to use:

The Double-Check Habit

Where does AI misinformation show up most in our classes/subjects?

What does "verify" look like in my subject area? (specific steps):

7) Practice Two Scenarios + Align

Scenario 1:

One-sentence safest student response (using shared language):

Scenario 2:

One-sentence safest student response (using shared language):

Shared wording we want everyone using (final):

Close: Immediate Classroom Action

Mini-lesson date (within 10 school days): _____

Delivery method: Whole group / Asynchronous / Other: _____

Exit ticket (final):

1. One new thing I learned...
2. One question I still have... Optional: Would you share this? Why or why not?

Next Meeting: Bring Evidence

What we'll bring: 3 anonymous student exit tickets (Strong / Developing / Needs support)

Patterns we'll look for:

- Identifying information awareness
- Billboard Test use
- Permission choices
- Verify-before-sharing habit
- How students describe AI (tool vs. relationship)

Closing reflection: One shared phrase I'll use this week + where I'll use it first: