EdSutra

Teaching guide

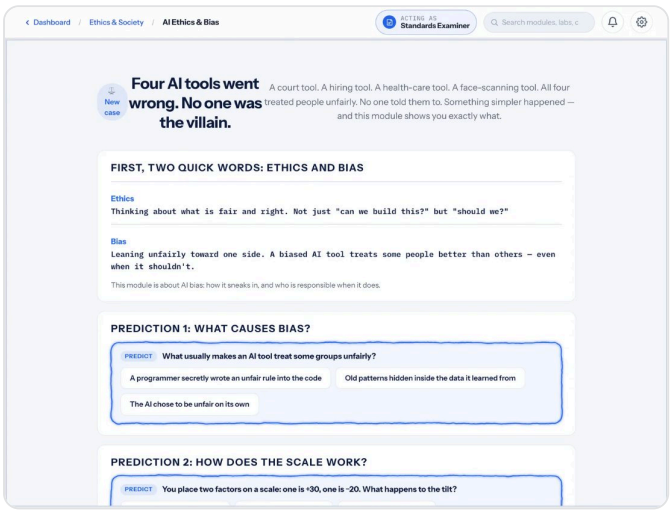
Module 05 · Ethics & Society · Unit 01 · Stage 3, Responsibility

AI Ethics & Bias

The Fairness Scale

Grades 8–12 | 50 minutes | Role: **Standards Examiner**

Grade 8: About age 13–14 · Year 9 in the UK · Class 8 in India



The big idea

AI bias rarely starts with someone writing an unfair rule. It builds up from ordinary choices: what data was collected, what stood in for gender or race, what the model was built to predict, and where it was used. Checks such as audits push back, but no check catches everything, and none changes the data a model learned from.

The learner's job

Learners are the **Standards Examiner** at Northvale Civic Standards Office. Northvale wants to buy an AI system. Nobody has asked where it could go wrong. Their task: find where bias gets into a system. Pick the checks that push back.

By the end, learners can say

- I can recall that bias enters through data, algorithm design, and deployment — not just intent
- I can predict which category caused a real documented AI-bias case, then check against the record
- I can operate the fairness scale and read the tilt as a real sum, not a mood ring
- I can defend which mitigations would cut a given bias fastest, citing case evidence

Before class

- One computer or tablet for each learner or pair, with EdSutra open (this Module is part of the full EdSutra studio).
- Open the studio once on each device before class, so the Unit still runs if the internet drops.
- Play the intro and the first case yourself, so you know where learners will click.
- Draw three columns on the board: Data, Algorithm, Deployment. You will sort the four cases into them during Analyze.
- Best after What is AI? and How Machines Learn.
- Units open in order, and Module 1 comes first. To start a class here, switch on Open every Unit in Settings → Teachers & parents.

Words to introduce

Ethics: Thinking about what is fair and right: not just can we, but should we.

Bias: Leaning unfairly toward one side, so some people are treated worse than others.

Proxy variable: A stand-in, like a word or zip code, that tracks gender or race.

Historical bias: Old unfair decisions inside the data, which a model learns and repeats.

Feedback loop: A model's outputs become its next data, so a pattern keeps growing.

Mitigation: A check that pushes back against bias, such as an independent audit.

Running the lesson 50 minutes

Open 5 min

"Four AI tools treated people unfairly: a court tool, a hiring tool, a health-care tool and a face scanner. Nobody told them to. So where did the unfairness come from?"

Before anyone clicks, ask: for each of the four tools, did the unfairness start in the data, in how the tool was built, or in where it was used? Write the guesses on the board and come back to them in Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Read the two quick words, ethics and bias, then answer the first prediction: what usually makes an AI tool treat some groups unfairly?	• In your own words, what does bias mean for an AI tool? • Why is "a programmer wrote an unfair rule" the wrong answer here?	Bias usually comes from old patterns in the data a tool learned from, not from someone meaning harm.
2. Understand 4 min	Answer the second prediction: one factor is +30, one is -20, so what does the scale show? Then step up to the Fairness Scale.	• What does the tilt number actually measure? • Why is a scale a useful picture of bias, and where does it fall short?	The tilt is a running total: +30 and -20 make +10. The weights are a ranking written for teaching, not measurements.
3. Apply 10 min	Click bias factors and checks to place them on the scale or take them off, and watch the tilt. Learners need at least four placed before their defense counts.	• Which single factor pushes the scale furthest toward biased? • What happens when you add an independent audit? Is the system now fair?	Checks pull the tilt back, but a balanced reading is not proof of fairness. That has to be checked on real results.
4. Analyze 15 min	Open the four real cases. For each, predict whether data, the algorithm or deployment was mainly responsible, then read the documented outcome and its lesson.	• The hiring tool never saw a gender column. Why did it mark down resumes with the word "women's"? • The health-care tool's cost records were accurate. Why was it still unfair? • Look at our guesses on the board. Which did we get wrong?	A word stood in for gender, and predicting cost carried old unequal spending. Bias entered through ordinary choices.
5. Evaluate 12 min	In DEFEND, pick the two checks to use first, choose the reason that backs them, and say what the plan cannot do alone. Then take the five-question quick check.	• Which two checks did you choose, and which case backs your reason? • What can audits and diverse teams not fix on their own?	Checks can find and flag bias, but the data or decisions the model learned from still have to change.

Steps follow Bloom's revised taxonomy, in the order the Unit presents them.

Deepen the thinking

Critical thinking	The scale's weights were written for teaching, not measured. When is a made-up model still useful, and when could it mislead?
AI thinking	The hiring tool never saw gender, yet found a stand-in. If you removed that word too, what new stand-in might it find?
Ethics and fairness	A town knows a face scanner works worse for some faces. Is using it anyway a technical problem or a decision someone made?
Connect to life	Think of a rule at school or online that seems neutral. Could it still work worse for some people? How?

Misconceptions to listen for

"AI bias means someone programmed it to be unfair."
None of the four cases began with an unfair rule. The causes were the data, what a tool was built to predict, and where it was used.

"Removing race or gender from the data fixes it."
The hiring tool never saw gender, yet it marked down the word "women's". A stand-in can carry the same pattern.

"If accuracy is high, the system is fine."
The Misaligned Objective factor and the cases show a tool can be accurate overall and still fail one group badly.

In real life the card learners see at the end of the Unit

- Hospitals, banks and job sites use scoring tools; old patterns in their data can treat some groups unfairly.
- Hiring tools in the news: one marked down applications saying "women's", though gender was never an input.
- Face-recognition tools can work worse for some faces when their training photos lacked people like them.

Try it this week. When a tool seems to work better for some people, guess where bias could enter: the data, the design, or how it's used.

Watch out for. AI bias rarely starts with someone writing an unfair rule. It usually comes from patterns in the data, and high accuracy can hide it.

Talk about it. Is a tool fair if it's right most of the time but wrong more often for one group?

Exit ticket

1. Name the three places bias can enter an AI system.

   _____

2. What is a proxy variable? Give the example from the hiring case.

   _____

3. Name one check that pushes back on bias, and one thing it cannot fix alone.

   _____

Going further

In small groups, learners pick one of the four cases and present a one-page plan: where the bias entered, the two checks they would add first, and what those checks still could not fix.

EdSutra

Teaching guide

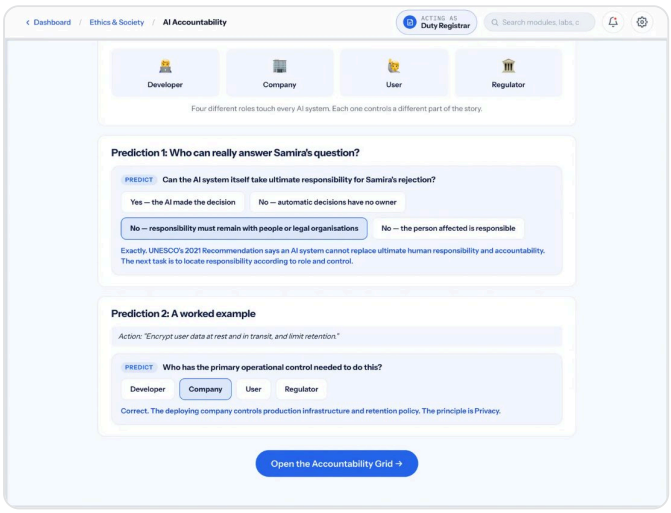
Module 05 · Ethics & Society · Unit 02 · Stage 3, Responsibility

AI Accountability

The Accountability Grid

Grades 7–12 | **45 minutes** | Role: **Duty Registrar**

Grade 7: About age 12–13 · Year 8 in the UK · Class 7 in India



The big idea

When an AI system causes harm, the system itself cannot answer for it. Responsibility stays with people and organizations, according to what each one controls: developers build and test, companies deploy and monitor, users decide how far to rely on it, and regulators set and enforce rules. Each duty has a primary home, but the work is shared.

The learner's job

Learners are the **Duty Registrar** at Northvale Register of Responsibility. Something will go wrong eventually. Today, everyone can still point at someone else. Their task: put a name against every duty. Before anything goes wrong.

By the end, learners can say

- I can differentiate the responsibilities of developers, companies, users, and regulators
- I can apply transparency, fairness, privacy, and safety principles to specific stakeholder actions
- I can identify distractor 'responsibility-washing' actions that deflect rather than allocate accountability
- I can defend a complete accountability map with reasoning grounded in stakeholder capability

Before class

- One computer or tablet for each learner or pair, with EdSutra open (this Module is part of the full EdSutra studio).
- Open the studio once on each device before class, so the Unit still runs if the internet drops.
- Play the intro and the Capability Detective yourself, so you know where learners will click.
- Draw the grid on the board: Developer, Company, User, Regulator down the side; Transparency, Fairness, Privacy, Safety across the top.
- Best after What is AI? and AI Ethics & Bias.
- Units open in order, and Module 1 comes first. To start a class here, switch on Open every Unit in Settings → Teachers & parents.

Words to introduce

Accountability: Being answerable for a harm: explaining it, fixing it and making it right.

Stakeholder: A role with a part in an AI system: developer, company, user, regulator.

Capability: What a role can actually do, like edit training data or stop a service.

Principle: One of the four lenses: transparency, fairness, privacy or safety.

Primary home: The role with the most direct capability to carry out an action.

Responsibility-washing: An action that looks like accountability but pushes blame away instead.

Running the lesson 45 minutes

Open 5 min

"When an AI system makes an unfair decision, everyone can point at someone else. So who really answers for it?"

Read Samira's story aloud from the intro, then ask: who is responsible for fixing it: the software, the school, the vendor, the developers, or nobody? Tally the votes and return to them in Evaluate.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Read Samira's story, follow the life cycle (build, deploy, use, govern) and meet the four roles: developer, company, user and regulator.	• What are the four roles that touch every AI system? • Which role sets the rules, and which runs the live service?	Developers build, companies deploy, users rely on the output, and regulators set and enforce the rules.
2. Understand 4 min	Answer two predictions: can the AI itself take responsibility for Samira's rejection, and who controls encrypting user data and limiting how long it is kept?	• Why can't the AI system be the one responsible? • Why is encrypting data the company's job, not the developer's?	An AI cannot investigate, fix or make amends. The company runs the servers, so it controls storage and how long data stays.
3. Apply 8 min	In the Capability Detective, pick who has the real capability in four short cases. In the Principle Lens, sort four actions by transparency, fairness, privacy or safety. First answers lock.	• Who decides whether a medical record goes into a prompt? • Comparing results across groups: which principle is that, and why?	The user controls what they type in. Comparing outcomes across groups checks fairness, even if privacy also matters.
4. Analyze 15 min	Solve the grid: place 16 action cards in their one right cell and drop 4 responsibility-washing cards in the Reject tray. Then open Review to read who else shares each duty.	• The card never names who does it. What clue told you? • Why does putting all the blame on the user belong in the Reject tray? • Look at the Review: does a primary home mean only that role is involved?	Read what the action needs, then find who can do it. Rejects push blame away instead of taking on a duty.
5. Evaluate 9 min	In DEFEND, pick the placement that was hardest, say whether who or which principle made it hard, pick the clue that settles it, and say why naming who matters. Then take the quick check.	• Which placement was hardest for you, and what settled it? • Look back at our votes on Samira's case. Would you change yours?	Only people and organizations can investigate and remedy a harm. Naming who acts first shows where to start.

Steps follow Bloom's revised taxonomy, in the order the Unit presents them.

Deepen the thinking

Critical thinking	The grid gives each action one primary home. What does that make easier to see, and what does it hide?
AI thinking	A company says its model's behavior just appeared and nobody controls it. What could it still test, watch or limit?
Ethics and fairness	Samira was told the decision was automatic. Is it fair to give no way to appeal? Who should create one?
Connect to life	Think of an app you use. Which of its duties are yours as the user, and which are not?

Misconceptions to listen for

"The AI is responsible because it made the output."

An output cannot investigate harm or offer a remedy. Prediction 1 shows responsibility stays with people and organizations.

"The user is always the one to blame."

Users have real duties, like checking high-stakes answers, but the grid also gives duties to developers, companies and regulators.

"A primary home means only that role is involved."

Review lists who else shares every duty. The grid shows who acts first, not who acts alone.

In real life the card learners see at the end of the Unit

- When a school or company says "the software decided", people still built, chose and ran that software.
- Apps often say "check important answers". That's a real user duty, but users aren't to blame for everything.
- Regulators can investigate when AI causes harm, using rules like data-protection laws, and order fixes.

Try it this week. With a parent or teacher, pick one AI tool you both know. Name who built it, who runs it, who uses it, and who checks on it.

Watch out for. An AI output can't be responsible: it can't fix harm or answer for it. Responsibility is shared among the people and organizations behind it.

Talk about it. If an AI tool made an unfair decision about you, who would you want to answer for it?

Exit ticket

1. Name the four roles in the Accountability Grid.

2. Why can't an AI system itself be held responsible?

3. What is responsibility-washing? Give one example from the Reject tray.

Going further

As a class project, groups pick an AI tool used at school or at home and list one real duty for each of the four roles, plus one responsibility-washing excuse they have heard and why it fails.

EdSutra

Teaching guide

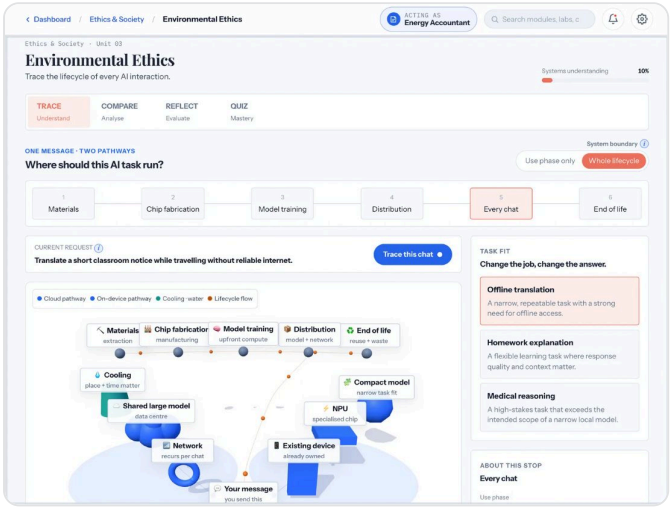
Module 05 · Ethics & Society · Unit 03 · Stage 3, Responsibility

Environmental Ethics

One Message · Two Pathways

Grades 7–12 | **40 minutes** | Role: **Energy Accountant**

Grade 7: About age 12–13 · Year 8 in the UK · Class 7 in India



The big idea

One AI message depends on a whole system: materials, chips, training, networks, data centers, cooling, devices and their disposal. There is no single cost per chat; the answer changes with the task and with what you count. The responsible choice is the least resource-heavy method that can do the task responsibly, with the whole lifecycle in view.

The learner's job

Learners are the **Energy Accountant** at Northvale Public Works · Energy Desk. Everyone asks what one message costs. The honest answer is: it depends on the road it takes. Their task: trace where one message's costs fall. Then say which road fits the job.

By the end, learners can say

- I can explain why an AI chat belongs to a material and technical lifecycle, not only a screen
- I can distinguish recurring use-phase demand from embodied hardware and upstream impacts
- I can compare cloud and on-device pathways across energy, network, cooling, hardware, and task capability
- I can apply proportionality by choosing a responsible pathway without hiding lifecycle trade-offs

Before class

- One computer or tablet for each learner or pair, with EdSutra open (this Module is part of the full EdSutra studio).
- Open the studio once on each device before class, so the Unit still runs if the internet drops.
- Play the intro and press Trace this chat once yourself, so you know where learners will click.
- Write three headings on the board: Before use, During use, After use. Learners sort the lifecycle stops into them in Understand.
- Best after What is AI? and AI Ethics & Bias.
- Units open in order, and Module 1 comes first. To start a class here, switch on Open every Unit in Settings → Teachers & parents.

Words to introduce

Environmental ethics: Judging an AI choice by its effect on materials, energy, nature and communities.

Lifecycle: Every stage of a system, from mined materials to disposal or recycling.

System boundary: What you choose to count: only the use phase, or the whole lifecycle.

Inference: The model working out an answer to one request.

On-device AI: A compact model that runs on your own phone or laptop.

Embodied hardware: The impact already built into hardware by mining and manufacturing it.

Proportionality: Using no more than the task really needs.

Running the lesson 40 minutes

Open 4 min

"Everyone asks what one AI message costs the planet. Why might the honest answer be "it depends on the road it takes"?"

Before anyone clicks, ask: for a quick translation and for advice about worrying symptoms, should the AI run on your phone or in the cloud? Note the votes and come back to them in Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 5 min	Read what environmental ethics means, then answer two predictions: where does the work of one AI message happen, and does running AI on your device make its impact disappear?	• Where does the work of one AI message really happen? • What is left when AI runs on your own device?	Across a whole connected system. On-device AI cuts some network and cloud work, but the device's own impacts remain.
2. Understand 11 min	On the systems map, press Trace this chat and follow the dots along the cloud and on-device paths. Open at least six stops, such as Materials, Cooling and End of life.	• Which stops happen before anyone sends a message? • Which path repeats a network trip for every request?	Materials, chips and training come before use. The cloud path makes a network trip every time; a local model does not.
3. Analyze 12 min	In Compare, switch between the three tasks and between Use phase only and Whole lifecycle. Learners compare the five separate bars for cloud and on-device each time.	• What appears when you widen the boundary to the whole lifecycle? • For medical reasoning, why does task fit matter more than the energy bar? • Look back at our votes. Did the task change your answer?	Hardware impacts appear. The phone suits translation, but a high-stakes task needs capability, and AI never replaces a doctor.
4. Evaluate 8 min	In Reflect, pick which of four rules best applies environmental ethics. Then take the five-question quiz; four right is a pass.	• Why is "always choose on-device" not good enough? • Why not just compare one carbon-per-chat number?	Use the least resource-heavy method that can do the task responsibly, and keep the whole lifecycle in view.

Steps follow Bloom's revised taxonomy, in the order the Unit presents them.

Deepen the thinking

Critical thinking	The Unit's bars are relative teaching estimates, not measured numbers. What would you need to know before trusting an exact claim about AI's cost?
AI thinking	A compact model fits on a phone but handles fewer kinds of task. Which of your own AI uses could it handle?
Ethics and fairness	Mining for AI hardware has local effects on the people who live nearby. Whose job is it to count those costs?
Connect to life	Think of a device at home that still works. What would help it stay useful for longer?

Misconceptions to listen for

"A chat happens only inside the app."
The map traces the network, data center, electricity, cooling and hardware behind every reply, and the stages before and after use.

"On-device AI is free for the planet."
It can skip the network trip for narrow tasks, but charging, manufacturing, updates and disposal remain. Widen the boundary to see them.

"There is one true cost for every chat."
The cost changes with the model, hardware, place, time, task and boundary. That is why the Unit shows separate bars, not one score.

In real life the card learners see at the end of the Unit

- A chatbot question can travel through networks to data centers that use electricity, cooling and hardware.
- Some phones run AI on the device itself, skipping the network trip but still using battery and hardware.
- Old phones and laptops in an e-waste bin are part of the story too: materials, manufacturing and end of life.

Try it this week. Before using AI for a small task, ask: could a simpler tool do it, like a dictionary, a calculator or a friend?

Watch out for. There's no single "cost of a chat". It depends on the model, hardware, electricity and what you count, and on-device AI isn't free either.

Talk about it. Should apps show the energy cost of using AI? What would you want to know?

Exit ticket

1. Name two lifecycle stages that happen before anyone sends a message.

2. What changes when you switch from use phase to whole lifecycle?

3. For offline translation, which path would you choose, and why?

Going further

For homework, learners list three AI tasks from their week and decide for each: cloud, on-device or no AI at all. In class, compare choices and the reasons behind them.

EdSutra

Teaching guide

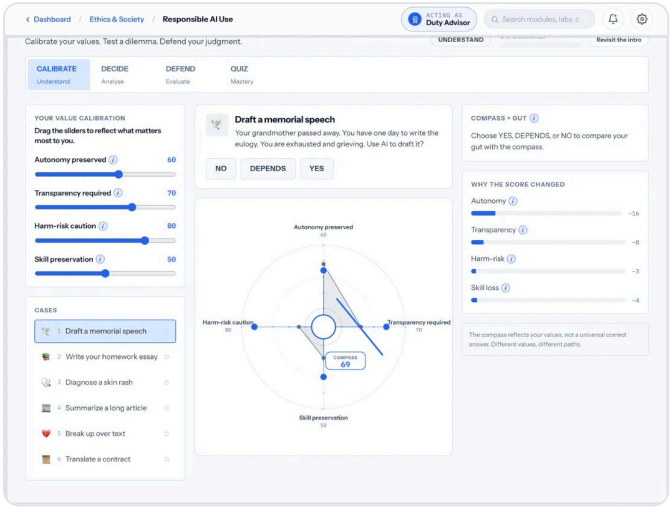
Module 05 · Ethics & Society · Unit 04 · Stage 3, Responsibility

Responsible AI Use

The Use-Case Compass

Grades 7–12 | **35 minutes** | Role: **Duty Advisor**

Grade 7: About age 12–13 · Year 8 in the UK · Class 7 in India



The big idea

Whether to use AI for a task is not one rule for everyone. It turns on what is at stake: whose work it is, who should know, what harm a mistake could cause, and which skills you would stop practicing. The Use-Case Compass mirrors your own values; you still own the decision and must be able to explain it.

The learner's job

Learners are the **Duty Advisor** at Northvale Advice Bureau. People bring real plans to this desk. They want permission. What they need is a way to think. Their task: help someone find their own answer. Make them able to defend it.

By the end, learners can say

- I can name the four value axes a use-case decision can be judged against
- I can explain why the same dilemma can produce different scores for different learners
- I can calibrate personal value weights and apply them to six real dilemmas
- I can compare a categorical gut judgment against a numeric alignment score and locate the exact source of disagreement
- I can defend a specific use-case judgment using the dilemma, the relevant value axis, and context

Before class

- One computer or tablet for each learner or pair, with EdSutra open (this Module is part of the full EdSutra studio).
- Open the studio once on each device before class, so the Unit still runs if the internet drops.
- Play the intro and decide one dilemma yourself, so you know where learners will click.
- No dilemma choice is marked wrong. Two cases touch on a death in the family and a breakup, so mention them first.
- Best after What is AI? and AI Accountability.
- Units open in order, and Module 1 comes first. To start a class here, switch on Open every Unit in Settings → Teachers & parents.

Words to introduce

Value weight: How much a value matters to you, set on a 0 to 100 slider.

Autonomy: Keeping your own choice, effort and authorship.

Transparency: People affected know AI was used and understand its limits.

Harm risk: How badly a wrong answer could hurt someone.

Skill preservation: Keeping practice at a skill that AI might do for you.

Compass score: Your weights combined with a dilemma's profile; higher means closer to your values.

Tension: When your gut answer and your compass score point different ways.

Running the lesson 37 minutes

Open 4 min

"People ask "Is it okay to use AI for this?" Why might the same question get different honest answers from different people?"

Before anyone clicks, ask: should you use AI to write a homework essay on a book you have not finished? Vote YES, DEPENDS or NO, and compare with the compass in Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Read the intro and answer two predictions: will two learners always get the same compass score, and what should happen when gut and compass disagree?	• Why can two learners get different scores for the same dilemma? • If your gut and the compass disagree, which one wins?	Their value weights differ. A disagreement is a reason to look closer, not a verdict to obey or ignore.
2. Understand 5 min	In Calibrate, move all four value sliders (autonomy, transparency, harm-risk caution, skill preservation) until they match what each learner values. Watch the compass score change.	• Which value did you set highest, and why? • Turn up one slider. Which dilemma's score changes most?	There is no correct setting. A higher weight matters most for dilemmas that put that value at stake.
3. Analyze 14 min	In Decide, read all six dilemmas, from a memorial speech to a lease, and answer YES, DEPENDS or NO. Then compare each answer with the compass and its per-value breakdown.	• Where did your gut and your compass pull in different directions? • Which value caused the biggest gap on that dilemma? • Look back at our homework vote. Did your compass agree?	Learners point to one value bar that explains the gap, and treat tension as something to explain, not an error.
4. Evaluate 10 min	In Defend, pick one decided dilemma: say how gut and compass relate, choose a reason that backs the decision, and name what would change it. Then take the quiz.	• Why does "nobody would know" never count as a reason? • What one condition would change your answer on this dilemma?	A reason tied to the case and their own decision, and a condition about a value the case really puts at stake.

Steps follow Bloom's revised taxonomy, in the order the Unit presents them.

Deepen the thinking

Critical thinking	The compass is a model someone wrote, with profiles someone chose. Which dilemma profile would you argue with, and why?
AI thinking	For the skin rash, AI might help you prepare questions for a doctor. Why is that different from trusting its diagnosis?
Ethics and fairness	If a friend sends you a heartfelt message, should they tell you AI helped write it? Does it depend on the message?
Connect to life	Think of a skill you are glad you learned the hard way. Would you have handed it to AI back then?

Misconceptions to listen for

"Each dilemma has one correct compass score."

The score multiplies your own weights by the dilemma's profile. Different weights give different scores for the same case.

"If my gut and the compass disagree, one is wrong."

The Unit never marks a choice wrong. Tension shows something worth explaining, and Defend asks learners to explain it.

"DEPENDS is a way to dodge the question."

DEPENDS is a real answer when you can name what it depends on. Defend asks for that condition.

In real life the card learners see at the end of the Unit

- Using AI to check spelling, brainstorm, or write a whole essay are three different choices, with different trade-offs.
- School AI rules can differ by subject and task: one tool can suit one job but not another.
- A birthday message written with AI help raises a question: should the person getting it be told?

Try it this week. Before using AI this week, ask: Is it still my work? Should people know? What if it's wrong? Will I lose a skill?

Watch out for. Whether you'd get caught isn't the point; the choice is the same either way. "It depends" counts when you can name what it depends on.

Talk about it. Which task would you never hand to AI, and which would you happily share with it?

Exit ticket

1. Name the four values the compass weighs.

2. What does it mean when your gut and your compass are in tension?

3. Why doesn't "I might get caught" count as a reason to change your answer?

Going further

As a class project, learners write their own dilemma about AI use at school, rate how much it puts each of the four values at stake, and swap with a partner to decide and defend.

EdSutra

Teaching guide

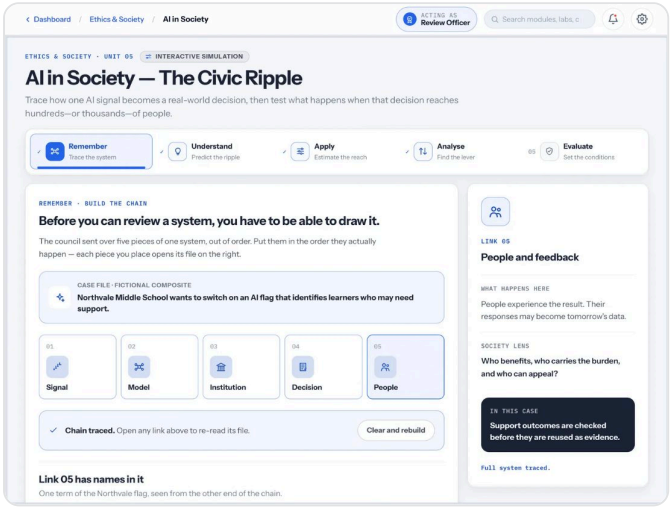
Module 05 · Ethics & Society · Unit 05 · Stage 3, Responsibility

AI in Society

The Civic Ripple

Grades 7–12 | 25 minutes | Role: **Independent Review Officer**

Grade 7: About age 12–13 · Year 8 in the UK · Class 7 in India



The big idea

An AI score changes lives only when an institution turns it into a decision that reaches people. Scale multiplies mistakes: a 6% miss rate across 1,000 decisions is 60 wrong calls. Human review helps only for the share of cases it covers, so a responsible launch matches its reach and safeguards to the stakes.

The learner's job

Learners are the **Independent Review Officer** at Northvale Civic Review Office. Three places want to turn on an AI system. Each one will make choices about real people. Their task: trace one AI decision from data to people. Make a finding you can defend.

By the end, learners can say

- I can identify the five links between an AI signal and a social outcome
- I can explain why institutions, operating rules, and feedback loops shape AI's impact
- I can calculate how reach, observed miss rate, and human review change unresolved misses
- I can compare deployment plans and locate which assumption creates the difference
- I can defend a launch posture with safeguards matched to the stakes

Before class

- One computer or tablet for each learner or pair, with EdSutra open (this Module is part of the full EdSutra studio).
- Open the studio once on each device before class, so the Unit still runs if the internet drops.
- Play the Remember step yourself, so you know where learners will click.
- Pairs work well: one learner drives, the other is the Review Officer who writes down each estimate and safeguard.
- Best after How Machines Learn and AI Ethics & Bias.
- Units open in order, and Module 1 comes first. To start a class here, switch on Open every Unit in Settings → Teachers & parents.

Words to introduce

Institution: An organization, like a school, that decides what an AI output means in practice.

Operating rule: The rule that turns a score into an action, such as counselor review.

Reach: How many decisions a system makes in one cycle.

Miss rate: The share of cases the system gets wrong, seen in a trial.

Review coverage: The share of possible misses that a person checks.

Feedback loop: A system's outputs shape the next round of its data.

Safeguard: A protection such as human review, an appeal path or outcome monitoring.

Running the lesson 29 minutes

Open 4 min

"A school's AI flag is right 94 times in 100. That sounds good. So why might the school still need to think hard before switching it on?"

Before anyone clicks, ask: if the flag makes 1,000 decisions a term, how many will it get wrong? And if a counselor reviews half of them, how many slip through? Write the guesses down for Apply.

Step	What learners do	Ask the class	Listen for
1. Remember 5 min	Put the five links of the school's support flag in order, from input signal to people. Each link opens its file. Then read about three learners at the last link.	• Who wrote the rule at the institution link, and who can change it? • Priya, Marcus and Amara had three different outcomes. What were they?	A correct flag, a wrong flag, and a missed learner who needed help. The model is only one link of five.
2. Understand 4 min	Make two predictions before the answer shows: what changed when a school swapped counselor review for automatic removal from an activity, and which part of a feed reinforces itself.	• The model's score stayed the same. Why did the effect on learners change? • How can a feed help create the pattern it claims to observe?	The institution attached a harsher action to the same score. Clicks on what is shown become tomorrow's training data.
3. Apply 6 min	Run the estimator: pick a reach (200, 1,000 or 5,000 decisions) and a review coverage, read the unresolved misses, and save two different settings to compare.	• At 1,000 decisions and no review, how many misses are left? • Why does even full review not bring the number to zero? • Look at our guesses from the start. How close were we?	60, since 1,000 times 6% is 60. Reviewers catch only 75% of what they see, so some misses remain.
4. Analyze 4 min	Compare two pairs of rollout plans side by side and name the input that explains the gap: reach, review coverage or the observed miss rate.	• The district rollout and the pilot share a 6% miss rate. What differs? • Why is one "impact score" less useful than three separate inputs?	Reach and review coverage changed together. Keeping inputs separate shows which lever made the difference.
5. Evaluate 6 min	For at least two of three cases (a school flag, a job screen, a library feed), choose launch, limited pilot or pause, and attach two or three safeguards that fit the stakes.	• Why can you pick at most three safeguards? • Which safeguard does the job screen need that the library feed does not?	Each safeguard costs time or money. High-stakes cases need human review or an appeal path; lower-stakes ones need monitoring or opt-out.

Steps follow Bloom's revised taxonomy, in the order the Unit presents them.

Deepen the thinking

Critical thinking	Amara was never flagged, so there was nothing to appeal. How could a school find the people its system never notices?
AI thinking	The flag reads attendance and grades. What could it never see, and how does that shape who it misses?
Ethics and fairness	Marcus spent a term explaining himself because nobody told him a record existed. What should people be told when AI scores them?
Connect to life	Where in your life does one decision reach many people at once, like a timetable or an app update?

Misconceptions to listen for

"A 94% success rate means the errors don't matter."

Apply multiplies the 6% miss rate by reach: 1,000 decisions leave 60 wrong calls before any review.

"Human review fixes every mistake."

Review covers only a share of cases and catches 75% of what it sees. Misses the flag never raised never reach it.

"AI's impact comes from the model alone."

The chain shows the institution's rule matters too: the same score caused more harm once the school attached a harsher action.

In real life the card learners see at the end of the Unit

- When a school or hospital uses an AI score, its rule decides whether a person reviews it first.
- A headline says an AI is "94% accurate". Across 1,000 cases, that's still 60 misses.
- Video apps reach millions of people, so even a tiny error rate can touch huge numbers.

Try it this week. Take a "95% right" claim from an ad or the news. Work out how many people the other 5% would be at your school.

Watch out for. A small error rate stops being small at scale, and human review only helps for the cases it actually covers.

Talk about it. Should a new AI tool start as a small pilot before everyone uses it? Why?

Exit ticket

1. Put the five links in order, from signal to people.

   _____

2. A tool with a 6% miss rate makes 5,000 decisions. How many misses?

   _____

3. Name one safeguard for a high-stakes decision, and say why it fits.

   _____

Going further

As a class project, groups design a pilot for an AI tool their school might use: how many people it reaches first, who reviews it, how people appeal, and what result would make them stop.

EdSutra

Teaching guide

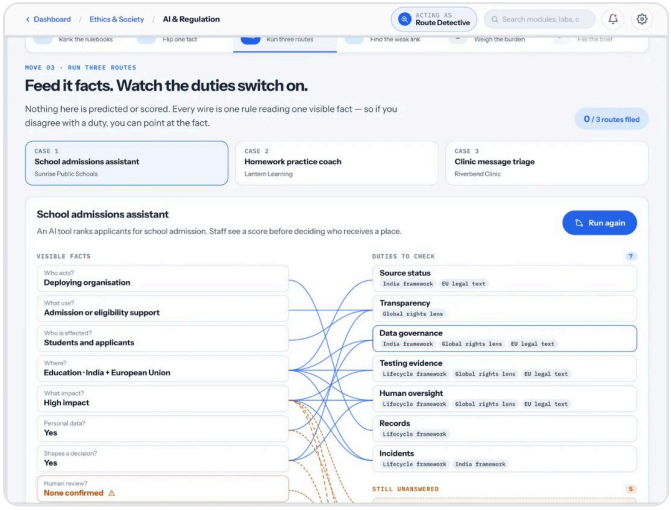
Module 05 · Ethics & Society · Unit 06 · Stage 3, Responsibility

AI & Regulation

The Rule Router

Grades 8–12 | 30 minutes | Role: **Governance Route Detective**

Grade 8: About age 13–14 · Year 9 in the UK · Class 8 in India



The big idea

There is no single worldwide AI rulebook. Laws, regulations, standards, policies and voluntary guidance carry different weight, and the same AI can follow different rules depending on its purpose, who it affects and where it is used. A route finder can point to duties and dated sources to check, but qualified people must confirm what applies today.

The learner's job

Learners are the **Governance Route Detective** at Northvale Civic Lab · Route Desk. The town has four AI plans. Each one crosses a different rule path. Their task: trace every turn. Check every source. Mark every doubt.

By the end, learners can say

- I can distinguish law, regulation, standard, policy, and voluntary guidance
- I can explain why the same AI capability can follow different governance routes in different contexts
- I can route three authored cases through inspectable conditions and attach relevant dated sources
- I can diagnose missing evidence across sectors and jurisdictions
- I can defend an oversight plan using duties, safeguards, costs, and uncertainty
- I can create a printable governance brief that records unresolved questions rather than claiming legal certainty

Before class

- One computer or tablet for each learner or pair, with EdSutra open (this Module is part of the full EdSutra studio).
- Open the studio once on each device before class, so the Unit still runs if the internet drops.
- Play the first two steps yourself, so you know where learners will click.
- Tell learners the cases are fictional and a route is not a legal conclusion, as the Unit says on screen.
- Best after AI Accountability and AI in Society.
- Units open in order, and Module 1 comes first. To start a class here, switch on Open every Unit in Settings → Teachers & parents.

Words to introduce

Law: A binding rule made by a legislature.

Regulation: Detailed binding rules made under a law's authority.

Standard: An agreed technical specification, binding only if a rule or contract says so.

Voluntary guidance: Recommended practice that is not, by itself, a law.

Deployer: The organization that puts an AI system to work in a real setting.

Governance route: The set of rules and duties a particular use of AI must check.

Recourse: A way for an affected person to question a result and get it fixed.

Running the lesson 32 minutes

Open 3 min

"A school policy, a national law and an expert's guidance are all called rules. Which can actually make someone stop?"

Before anyone clicks, ask: one AI ranks practice questions and the same kind of AI ranks school applicants. Should both follow the same rules? Note the votes and come back to them in Understand.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Rank the rulebooks: place law, regulation, standard, policy and voluntary guidance on the rung that matches their power. Then say what a school is when it buys and uses a ranking tool.	• Which kind of rule binds only the people inside one organization? • The school did not build the tool. Is it a provider or a deployer?	Law and regulation bind everyone they cover; a policy binds one organization. The school is the deployer.
2. Understand 4 min	Two pairs of tools share the same capability. Predict which side needs the deeper rules, then switch the board between them to see which facts changed.	• The model did not change. What made the admissions tool need more checks? • A clinician is "in the loop". When does that count as real oversight?	Its purpose and impact: it shapes who gets a school place. Oversight needs time, evidence and authority to change the result.
3. Apply 6 min	Run three cases through the router: admissions, homework coach and clinic triage. Each visible fact lights up duties; learners attach dated sources from the shelf and save each route.	• Pick one duty on the board. Which fact switched it on? • Why does the homework coach get a much shorter route?	Every duty traces back to a visible fact. Low impact and a teacher in between keep the coach's route short.
4. Analyze 4 min	Work through three weak-link cases, such as a clinic where a human is present but the review time is unknown, and pick the missing fact that matters most.	• Why can't the route board give the final legal answer? • What would show that a rejected job applicant has real recourse?	Coverage and status still need checking in the sources. Recourse needs notice, an appeal route and someone who can change the result.
5. Evaluate 5 min	For the admissions case, choose an oversight posture and safeguards, mark which duties they answer, name one real cost, and mark what is still unanswered.	• Which safeguard costs the most staff time, and is it worth it here? • Why can more oversight sometimes be worse, not better?	Any posture can be defended if it is tied to route evidence, names a real cost and leaves open questions visible.
6. Create 6 min	Build the Route Brief from choices: pick a case, state its purpose, name who carries it, whose rights are in play, the evidence and sources, open questions, a review date and a recommendation.	• Which purpose statement did you choose, and what does it put first? • What question did you leave open, and who should answer it?	A brief someone else can check: it names people, evidence and dated sources, and admits what is still unknown.

Steps follow Bloom's revised taxonomy, in the order the Unit presents them.

Deepen the thinking

Critical thinking	The router never says what the law requires, only what to check. Why might that be more honest than a yes-or-no answer?
AI thinking	The same ranking model can be a homework helper or an admissions gate. Why do rules follow the use, not the model?
Ethics and fairness	Every safeguard costs someone time, money or delay. Who should pay for keeping a high-impact AI safe?
Connect to life	Which rules at your school are school policy, and which come from a law? How could you find out?

Misconceptions to listen for

"There is one worldwide rulebook for AI."
The Unit's sources include a binding regulation, government guidelines, a voluntary framework and an international recommendation, each with its own reach.

"If the model is the same, the rules are the same."
In Understand, the same ranking model needs a deeper route once it shapes school admissions instead of suggesting practice questions.

"The route board tells you what the law requires."
It shows candidate sources and missing facts. Qualified people must still check the current text before anyone relies on it.

In real life the card learners see at the end of the Unit

- Age limits and privacy settings on apps exist partly because of laws protecting children's data.
- A school's rules on AI bind people in that school; a national law binds everyone it covers.
- The same AI tool can face different rules in different countries, or when used for hiring instead of games.

Try it this week. With a parent or teacher, look at one app's age rating or privacy settings. Is each one there because of a law, or company policy?

Watch out for. There's no single worldwide AI rulebook, and a route finder only points to rules to check. Experts must confirm what applies today.

Talk about it. What's one AI rule you think every country should have? Who would enforce it?

Exit ticket

1. Put these in order of power: policy, law, voluntary guidance.

   _____

2. Name one fact that can change an AI tool's governance route.

   _____

3. Why should a brief list the questions that are still open?

   _____

Going further

As a class project, learners pick one app they use and trace its route: who provides and deploys it, who it affects, and which rules might apply. Then they list what an expert would still need to check.