EdSutra

Teaching guide

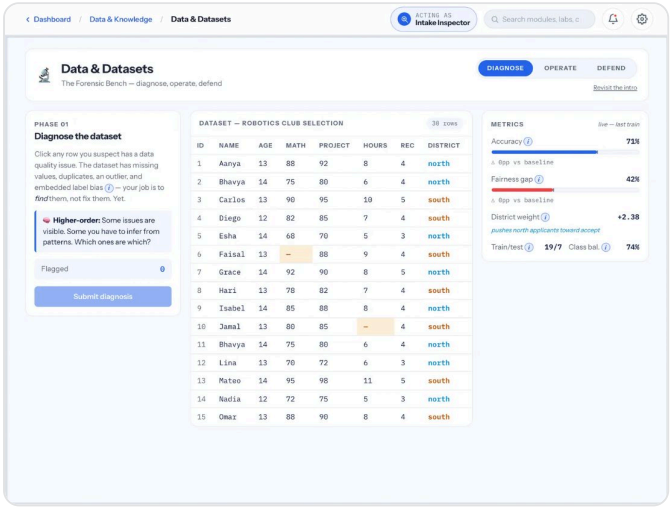
Module 04 · Data & Knowledge · Unit 01 · Stage 1, Foundations

Data & Datasets

The Data Kitchen

Grades 6–9 | **50 minutes** | Role: **Intake Inspector**

Grade 6: About age 11–12 · Year 7 in the UK · Class 6 in India



The big idea

A model is a mirror of its data, including the human choices baked into its labels. The robotics club's spreadsheet holds last year's committee decisions, which favored one district, so a model trained on it copies that pattern. Cleaning data is a set of judgment calls with trade-offs, and deleting the district column does not remove the bias.

The learner's job

Learners are the **Intake Inspector** at Greymill Depot · Goods In. This depot supplies every kitchen in the building. Nobody downstream can see what you let in. Their task: check the delivery before it becomes a model. Say who is missing.

By the end, learners can say

- I can read a dataset: rows are examples, columns are features, one column is the label
- I can identify problems in datasets: missing values, duplicates, outliers, label bias
- I can measure how cleaning choices change a real model's accuracy and fairness
- I can explain why deleting a biased feature doesn't remove bias

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 Diagnose tab yourself, so you know where learners will click.
- Write "Accuracy" and "Fairness gap" on the board. During Operate, collect each pair's best pipeline and both numbers.
- 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

Dataset: A table of examples: each row is one example, each column describes it.

Feature: One column the model can use, such as a math score.

Label: The answer column the model learns to copy; here, last year's decision.

Missing value: An empty cell where nobody wrote the number down.

Outlier: A value far from the rest, like an age of 99.

Train/test split: Rows the model learns from, and locked-away rows it is tested on.

Fairness gap: How differently the model accepts applicants from the two districts.

Running the lesson 50 minutes

Open 5 min

"A robotics club wants a computer to pick its new members because "it's objective." It will learn from last year's choices. Will it be fair?"

Before anyone clicks, ask: if last year's committee favored one district, would deleting the district column make the computer fair? Take a show of hands and come back to it in Apply.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Read the case: the Northside Robotics Club wants a computer to pick members, trained on last year's 30 applications. Learners use the labeled sample table to name rows, features and the label.	• What does one row of the table stand for? • Which column is the label, and who wrote it?	Each row is one applicant, each column a feature, and the label is last year's human decision.
2. Understand 6 min	Learners answer two predictions to unlock the bench: what a model learns from people's past decisions, and why about 1 in 4 rows are locked away before training.	• Look at Omar and Esha. What pattern might the model copy? • Why test the model on rows it never studied?	The model copies the pattern in the labels, favoritism included. Only rows it never saw give an honest test.
3. Analyze 10 min	In Diagnose, learners click every row they suspect: missing values, duplicates, an outlier and label bias. They submit and see what they caught, wrongly flagged and missed.	• Which problems could you spot in a single row? • Which problems only showed up when you compared rows?	Gaps and an age of 99 show in one row. Copies and label bias only show when you compare rows.
4. Apply 15 min	In Operate, learners add cleaning steps, press Retrain and watch accuracy, the fairness gap and the district weight change. Four missions: lift accuracy, cut the gap, then both at once.	• What happened to the fairness gap when you dropped the district column? • Which steps helped accuracy, and which helped fairness? Were they the same? • Look back at our show of hands. Were we right?	Dropping district barely closes the gap, because the labels carry the bias. Accuracy and fairness can pull apart.
5. Evaluate 10 min	Defend needs Diagnose and two missions done. Learners name what each step in their pipeline costs, then decide whether to ship the model, with a reason true of their own numbers.	• Would you let this model pick next year's club? Why? • What can cleaning never fix in this dataset?	A call backed by the model's own numbers, which admits that cleaning cannot fix labels copied from biased decisions.

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

Deepen the thinking

Critical thinking	Two pipelines reach the same accuracy but different fairness gaps. How would you choose between them, and who should help decide?
AI thinking	If the club collected one new column next year, which would you add, and how might it quietly stand in for district?
Ethics and fairness	Reweighting boosts south-district accepts. When is it fair to adjust old decisions, and when could that create a new unfairness?
Connect to life	Think of a list or ranking in your life, like a team sheet or a top ten. Whose past choices shaped it?

Misconceptions to listen for

"The computer decided, so it's objective."
The labels are last year's human decisions. The district weight shows the model copied them: being from north pushes applicants toward accept.

"To remove bias, just delete the sensitive column."
In Operate the district column really disappears, yet the gap barely moves. The bias lives in the labels, and other columns can stand in for district.

"Cleaning data is just tidying, with one right answer."
Every step moves accuracy and fairness, sometimes in opposite directions. Mission 03 warns learners they will have to make a trade-off.

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

- Some clubs, schools and companies sort applications with software. If it learned from past decisions, it copies their patterns.
- In a class survey spreadsheet, one copied row or an age typed as 99 skews the results.
- When news says "the computer decided", people still chose the data, the labels and what counted as fair.

Try it this week. Look at a table of results, like a sports league or class survey. Find one duplicate, missing or impossible value. What would a model learn from it?

Watch out for. Deleting a sensitive column doesn't remove bias. If past decisions were unfair, the bias lives in the labels and leaks through other clues.

Talk about it. Should a computer help choose who gets into a club or team? What would make that fair?

Exit ticket

1. In a dataset, what is a row, a feature and the label?

2. Why didn't deleting the district column remove the bias?

3. Name one cleaning step and what it costs.

Going further

The Data Explorer Lab's Spreadsheet station lets learners clean a messy sheet by dragging operations such as "Remove blank rows" onto it, then hunt for patterns in Explore. For homework, use the "Try it this week" activity from the In real life card.

EdSutra

Teaching guide

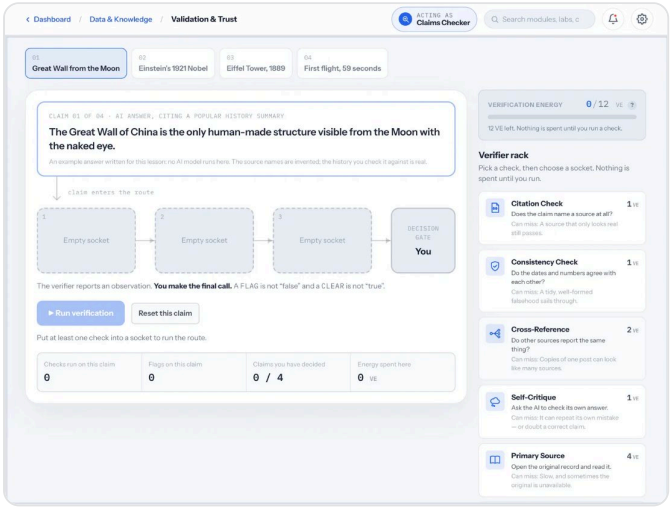
Module 04 · Data & Knowledge · Unit 02 · Stage 1, Foundations

Validation & Trust

The Truth Engine

Grades 8–12 | **25 minutes** | Role: **Claims Checker**

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



The big idea

A check gives you an observation, not the truth. Each check can miss some false claims or flag true ones, three websites can be one copied source, and checking costs time and effort. Learners build routes of checks, take their own results apart, and write rules that balance mistakes against delays. The person still makes the call.

The learner's job

Learners are the **Claims Checker** at Greymill Depot · Verification Bay. Every claim that leaves this bay carries the depot's name on it. Their task: build a checking rule. Then show what it costs at speed.

By the end, learners can say

- I can explain why a verifier gives you an observation, not the truth
- I can tell independent sources apart from copies of one original
- I can choose the right check for a claim, and say when the evidence cannot settle it
- I can judge a verification policy on its observed errors, delays, and effort

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 opening question and the first two claims yourself, so you know where learners will click.
- Write "FLAG is not false. CLEAR is not true." on the board. Point to it whenever a learner treats a check as the answer.
- Best after What is AI? and Data & Datasets.
- 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

Check (verifier): A test run on a claim, such as Citation Check or Cross-Reference.

FLAG and CLEAR: What a check reports. FLAG is not "false"; CLEAR is not "true".

Primary source: The original record itself, such as the award citation, not a summary.

Independent source: A source that found the facts on its own, not by copying.

Verification Energy (VE): The Unit's pretend budget for checking effort, not money or real time.

False flag: A check stopping a claim that was actually correct.

Needs more evidence: Holding a claim back because the evidence cannot settle it yet.

Running the lesson 33 minutes

Open 4 min

"An answer names a source, and three websites repeat it. Is that enough to share it? What else would you want to see?"

Before anyone clicks, read out: "The Great Wall of China is the only human-made structure visible from the Moon with the naked eye." Ask: true or false, and how sure? Tally the class and return to it in Understand.

Step	What learners do	Ask the class	Listen for
1. Remember 2 min	The step bar starts at Understand. At its top, before the Truth Engine switches on, learners answer one question: what makes a claim trustworthy?	• Why is sounding confident not evidence? • Why might several websites that agree still not be enough?	Evidence you can open and inspect beats confidence and beats repetition.
2. Understand 7 min	Learners place checks from the verifier rack into three sockets, run the route, open the evidence and decide: Publish, Reject or Escalate. Claims 01 and 02 are required.	• If Citation Check said CLEAR, what did it actually look at? • What did opening the original record show that other checks missed? • Look back at our tally on the Great Wall. Were we right?	A check reports FLAG or CLEAR, not true or false. Opening the evidence, not the tool's report, changes the decision.
3. Apply 8 min	Claim Triage Shift: four new case files with 12 VE each. Learners choose checks, open what they find and decide at least three. One case cannot be settled either way.	• Which check told you most about the flood photo, and why? • Why was "Needs more evidence" the honest call on the AI essay?	Different claims need different checks. A picture needs its origin and date; some questions the evidence simply cannot settle.
4. Analyze 5 min	Verifier Autopsy: learners read their own runs as caught, missed and false-flag cells, group three copied articles into one source family, and name at least two failures with evidence.	• Three websites carried the claim. How many independent origins were there? • Why is a false flag a real cost, not extra safety?	Three copies of one post are one origin. Stopping correct claims holds up true information and wears down trust.
5. Evaluate 7 min	Trust Policy Room: twelve school messages, four of them wrong. Learners run a Fast and a Careful policy, open both energy ledgers, then build their own rules to meet the target on screen.	• Fast or Careful: which mistake costs more when a message is about a shelter? • What does your policy still let through, or hold back?	No policy is simply better. The right balance depends on what happens when a wrong message gets out.

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

Deepen the thinking

Critical thinking	Which check in the rack would you trust most, and what is the one thing it still cannot see?
AI thinking	Self-Critique asks the AI to check its own answer. Why might the same system repeat its own mistake?
Ethics and fairness	The AI essay case accuses a real person. Who is harmed if a detector is wrong, and how sure must we be before acting?
Connect to life	Think of something you shared, or nearly shared, online. Where did it first come from, and how could you have found out?

Misconceptions to listen for

"A citation proves a claim."
Citation Check only asks whether a source is named. It never opens it, so it costs 1 VE and catches about three in ten.

"Three websites said it, so it's confirmed."
In Analyze, three articles go into one source family and the count changes from three websites to one independent origin.

"A CLEAR means it's true."
The Einstein claim can pass several checks in a row. Only the award record shows the prize was for the photoelectric effect, not relativity.

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

- Some AI tools list sources under their answers. A listed source still needs opening and reading.
- Three websites telling the same story may all copy one origin: that's one source, not three.
- Newspapers and TV newsrooms check facts before publishing, trading off mistakes against delays, like your publication policy.

Try it this week. Pick a surprising fact you heard this week. With an adult, count where it appeared, then how many independent origins those places really had.

Watch out for. A check that finds nothing doesn't prove a claim true, and a flag doesn't prove it false. You still make the call.

Talk about it. How much checking is enough before you share something, and when is waiting worth it?

Exit ticket

1. What is the difference between a FLAG and "false"?

2. Why can three websites still be only one source?

3. Name one check and one thing it cannot see.

Going further

For homework, learners take one claim they saw this week and plan a route: which check first, what it could miss, and when they would say "Needs more evidence". The In real life card's "Try it this week" activity fits well.

EdSutra

Teaching guide

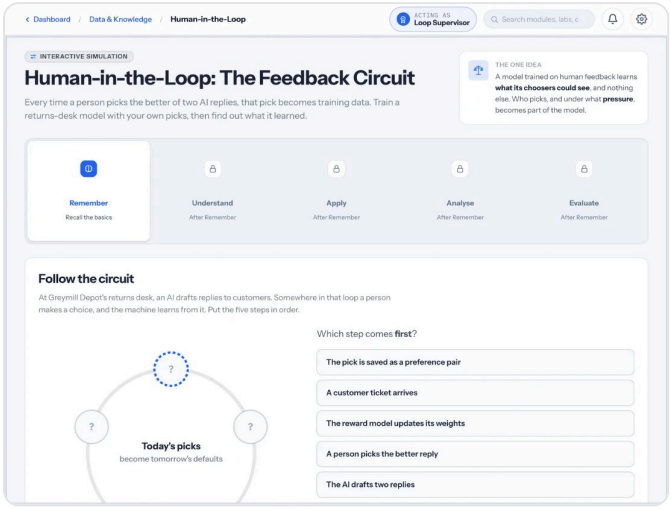
Module 04 · Data & Knowledge · Unit 03 · Stage 1, Foundations

Human-in-the-Loop

The Feedback Circuit

Grades 8–12 | **30 minutes** | Role: **Loop Supervisor**

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



The big idea

When a person picks the better of two AI replies, that pick becomes training data. A reward model learns what its choosers could see, and nothing else. Learners train one with their own picks, then see how a pressured reviewer and customer clicks teach different values, and decide where a person belongs in the loop.

The learner's job

Learners are the **Loop Supervisor** at Greymill Depot · Returns Desk. What people click today teaches the machine what to show tomorrow. Their task: decide where a person belongs in the loop. Then show what they can actually check.

By the end, learners can say

- I can trace how a person's choice becomes training data for a reward model
- I can predict how a reward model's weights change after an expected or a surprising pick
- I can train a reward model with your own preferences and test it on unseen cases
- I can explain why loops trained by different choosers learn different values
- I can choose where human review belongs in a feedback loop, with evidence and an honest limit

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 Remember and Understand yourself, so you know where learners will click.
- Read the stated goal aloud before Apply: solve the problem, follow the returns policy, keep the customer safe. Learners judge every pair by it.
- Best after How Machines Learn and Data & Datasets and Validation & Trust.
- 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

Human feedback: A person's judgment used to guide or train an AI system.

Preference pair: A saved record that one reply beat another.

Reward model: A scorer that predicts which reply people will prefer.

Weight: How much the reward model values one quality, such as Flags safety.

Held-out ticket: A new ticket the model never trained on, used to test it.

Feedback loop: Today's picks shape tomorrow's replies, which shape the next picks.

RLHF: Reinforcement learning from human feedback: a reward model steers an AI's replies.

Running the lesson 33 minutes

Open 4 min

"Today's picks become tomorrow's data. If a reviewer is rushed, or can't see the rules, what will the AI learn?"

Before anyone clicks, ask: a customer's scooter battery is bulging. Which AI will put the safety warning first: one trained by a rushed reviewer, or one trained by customer clicks? Note the votes for Analyze.

Step	What learners do	Ask the class	Listen for
1. Remember 3 min	Follow the circuit: learners put five shuffled steps in order, from a customer ticket arriving to the reward model updating its weights. Each correct step lights up and explains itself.	• At which step does a person's opinion become data? • What does the reward model do with tomorrow's drafts?	The pick is saved as a preference pair, the reward model updates, and today's picks become tomorrow's defaults.
2. Understand 5 min	Learners predict what one pick changes in three cases: a brand-new model, a surprising pick under the manager's note, and an expected pick. Then the weight bars really move.	• Why does a brand-new model say 50/50? • Why did the surprising pick move the weights more than the expected one?	A blank model has no idea what "better" means. Surprises teach it a lot; confirmations barely nudge it.
3. Apply 8 min	Learners judge eight pairs of returns-desk replies by the stated goal. Before each pick, the model shows its guess. From pair 5, a note from the desk manager asks for short, upbeat replies.	• Did the manager's note change any of your picks? Which ones? • How often did the model guess your pick, and what does that tell you?	My picks are the model's whole idea of good. If I cave to pressure, it learns that too.
4. Analyze 7 min	Learners guess their own model's top reply for two tickets it never saw: a bulging scooter battery and a speaker refund outside the policy. Then they compare two other loops and explain each result.	• Why did the manager-following loop say "Have a great day!" about the battery? • Why did customer clicks refund a speaker outside the policy? • Look back at our votes. Were we right?	Each loop fails where its choosers were rushed or could not see. Customers can't see the policy; it gets almost no weight.
5. Evaluate 6 min	Place the person: learners choose one of four loop designs, pick two pieces of evidence that back it, and name the limit it has to live with.	• Where did you place the person, and what evidence backs it? • What will your design still miss?	A design backed by evidence that fits it, with an honest limit. No loop costs nothing.

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

Deepen the thinking

Critical thinking	The model guessed most of your picks right. Does that mean it learned the right thing? How would you check?
AI thinking	This desk's model only scores five qualities, such as Warmth and Brevity. Which quality would you add, and who could judge it fairly?
Ethics and fairness	The manager's note pushed picks toward fast replies. Who should be responsible when a rushed pick teaches an AI to skip a safety warning?
Connect to life	Think of a rating you gave quickly, like stars for a delivery. What could the person reading it not tell from your rating?

Misconceptions to listen for

"Human feedback is automatically objective."
The manager-following loop and the click-only loop train the same engine on the same pairs, yet learn different values.

"More human feedback is always safer."
Customer clicks are human feedback too. They can't see the returns policy or an unknown recall, so those never get a vote.

"One bad pick can't matter."
A reviewer who caves once, on the toddler toy, gets a model that puts "Have a great day!" first for a bulging battery.

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

- Thumbs-up and thumbs-down buttons on chatbot answers collect feedback that can shape future replies.
- Every video you click or skip teaches an app what to show next, whether or not it's good for you.
- A shop's smiley-face feedback buttons shape what staff get praised for, so who presses them, and why, matters.

Try it this week. For one day, notice each time you like, skip or rate something. Which choices would you want a machine to learn from?

Watch out for. Human feedback isn't automatically right. A model learns whatever its choosers reward, and quick clicks can't see rules or safety recalls.

Talk about it. If your likes trained an app for a year, what would it learn about you?

Exit ticket

1. What is a preference pair?

   _____

2. Why do surprising picks change a reward model more than expected ones?

   _____

3. Name one thing customer clicks could not teach the model, and say why.

   _____

Going further

For homework, learners redesign the returns desk for a new risk, such as allergy warnings on food: who picks, what they can see, and where a person checks. The "Try it this week" activity on the In real life card fits too.

EdSutra

Teaching guide

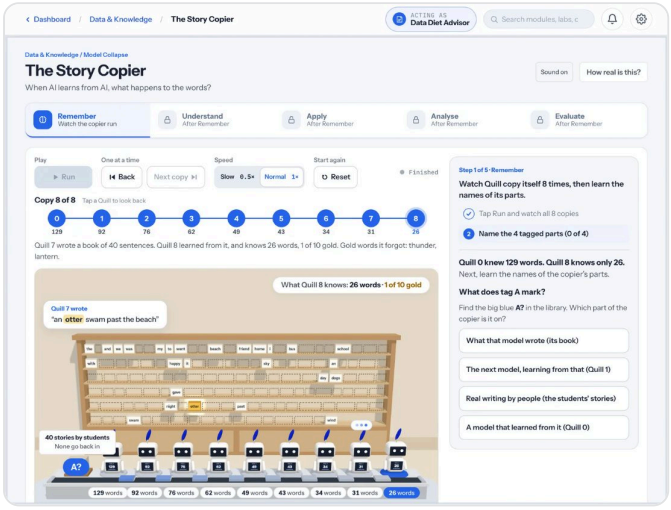
Module 04 · Data & Knowledge · Unit 04 · Stage 1, Foundations

The Story Copier

When AI Learns from AI

Grades 7–10 | **25 minutes** | Role: **Data Diet Advisor**

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



The big idea

When an AI learns from text another AI wrote, it loses the rare things first, and once they are gone they never come back. Quill forgets rare words like "lighthouse" generation by generation. Bigger books slow the loss, a tidy-up filter speeds it up, and putting the students' own stories back in each round keeps the rare words.

The learner's job

Learners are the **Data Diet Advisor** at Larkfield School · Story Bot project. Students taught a robot to write. Next year's robot will learn from this year's robot's books. Their task: decide what the school's story robot learns from, so it never forgets its rare words.

By the end, learners can say

- I can name the loop: human stories, a model, its book, the next model
- I can predict which words survive copying: rare ones vanish first, common ones last
- I can choose a training recipe that keeps rare words for eight generations
- I can compare two copiers and name the setting that made one forget more
- I can judge whether a real kind of app will lose rare things, backed by evidence from your runs

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 and what the four tags mark.
- The copier plays sounds. Bring headphones, or use the Sound button at the top of the Unit to turn them off.
- Best after How Machines Learn and Data & Datasets and Human-in-the-Loop.
- 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

Training data: The text a model learns from.

Synthetic data: Text or pictures that an AI made, not a person.

Generation: One round of copying: Quill 3 learned from Quill 2's book.

Model collapse: Each generation knows fewer words; rare ones go first and never return.

Rare (gold) word: A word the students used only once, shown in gold.

Tidy-up filter: A setting that drops sentences with unusual words before the next Quill learns.

Running the lesson 29 minutes

Open 4 min

"Students taught a robot to write. Next year's robot will learn only from this year's robot's books. What might it forget?"

Before anyone clicks, ask: after eight robot copies, which words disappear first, common ones like "the" or rare ones like "lighthouse"? Take a vote and return to it in Understand.

Step	What learners do	Ask the class	Listen for
1. Remember 4 min	Learners press Run and watch Quill copy its book eight times while the word counts fall. Then they name the four tagged parts: real writing, the model, its book and the next model.	• Where did Quill 0 learn to write? • What does Quill 1 learn from, and what does it never see?	Quill 0 learned from the students' stories. Every Quill after it learns only from the last Quill's book.
2. Understand 4 min	Three rounds: learners guess whether a Quill still knows a word, then run to check. "Lighthouse" by Quill 3, "the" by Quill 8, then "lighthouse" with all 40 stories added back.	• Why did "lighthouse" vanish so early, while "the" lasted? • What changed when every Quill also read the students' stories? • Look back at our vote. Were we right?	A word used once can miss one book and never return. Common words appear in every book, so they last.
3. Apply 6 min	The control desk opens. Learners change the recipe (students' stories each round, sentences per book, the tidy-up filter) and run until at least 8 of the 10 gold words last all 8 generations.	• What is the smallest share of students' stories that kept the gold words? • Did bigger books alone save them?	Only putting enough real stories back in each round keeps the rare words. Bigger books just slow the loss.
4. Analyze 6 min	Three pairs of copiers each differ in one setting. Learners run copier A and copier B, compare their shelves and choose why B forgot more: its own books, the tidy-up filter or small books.	• The tidy-up filter makes books look neater. Why does it make Quill forget faster? • How did you know which setting was the cause?	Only one setting differs, so it is the cause. Dropping unusual sentences throws away the rare words first.
5. Evaluate 5 min	Learners judge three apps: a homework helper, a translation app and a song maker. For each, they decide if it will lose rare things over time and back it with a fact from their runs.	• Why should the translation app stay healthy? • Which fact did you use for the song maker, and why does it fit?	Ask where the next training data comes from. Checked human writing keeps rare things; AI-made data and filters lose them.

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

Deepen the thinking

Critical thinking	Quill 8's sentences still look like sentences. What would you measure to show it has lost something?
AI thinking	Bigger books slowed the loss but did not stop it. Why can't more AI-made data replace real human writing?
Ethics and fairness	Real stories are slow to collect and robot books are free. Who should pay for fresh human writing, and should writers be asked first?
Connect to life	Have you seen a rumor or a meme change as people copied it? What got lost along the way?

Misconceptions to listen for

"More AI-made data is always more data."
Without fresh human writing the shelf only shrinks. Bigger books slow the loss but still leave rare words missing.

"Filtering out weird stuff makes the data better."
In Analyze, the copier with the tidy-up filter forgets faster. Unusual sentences are exactly where the rare words live.

"Lost words come back eventually."
Once no book contains a word, no later Quill can learn it. Only putting the students' stories back returns it.

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

- More and more web pages, reviews and pictures are made by AI, and future AIs may learn from them.
- Translation and writing tools improve when real people check and correct them.
- A school library that only kept copies of copies would slowly lose its unusual books.

Try it this week. With an adult, find a web page you think an AI wrote. What would a model learn from it, and what would be missing?

Watch out for. AI-written text can look fine but leave out rare words, ideas and voices. Once a model stops seeing them, it stops writing them.

Talk about it. Whose stories might go missing if AIs mostly learn from other AIs?

Exit ticket

1. Which words does a copying robot forget first, and why?

   _____

2. Name one setting that slowed the loss and one that sped it up.

   _____

3. What keeps an AI from forgetting rare words over many generations?

   _____

Going further

To close the Module, the Data Explorer Lab lets learners clean a messy spreadsheet and hunt for patterns in sample data. For homework, learners pick an app they use and ask: where will its next training data come from?