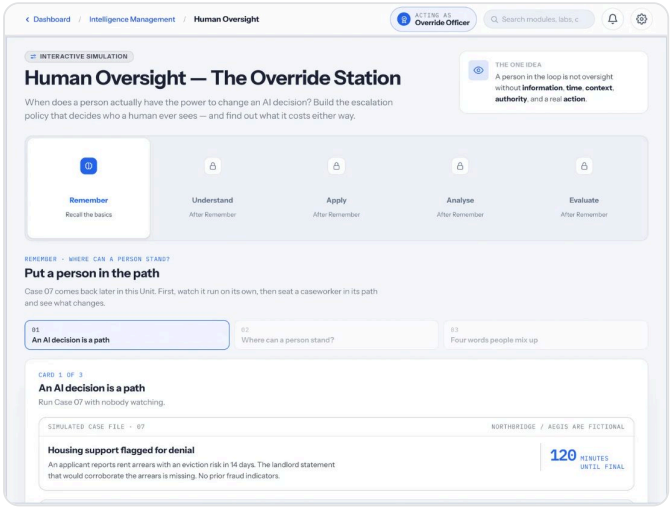
Module 09 · Intelligence Management · Unit 02 · Stage 3, Responsibility

Human Oversight

The Override Station

Grades 8–12 | 25 minutes | Role: **Override Officer**

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



The big idea

A person in the decision path is real oversight only if they have the information, time, context, authority and a real action to change the outcome. Sending too few cases to people misses harm; sending too many makes urgent cases wait. A good escalation policy is judged on its counts: missed, late and unnecessary.

The learner's job

Learners are the **Override Officer** at Ironvale Control · Override Station. Every case here has a deadline. Review takes time you do not have. Their task: decide what a human must see first. Live with what you let past.

By the end, learners can say

- I can distinguish human in, on, and out of the loop — and review, escalation, override, and appeal
- I can judge whether a reviewer has the information, time, context, authority, and recourse to change an outcome
- I can configure an escalation policy and route a full case stream through a review queue
- I can trace a missed case, a late review, and a rubber-stamp back to the rule or interface that caused it
- I can defend a policy on its exact missed, late, and unnecessary counts rather than on intention

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.
- Draw three columns on the board, "Missed", "Late" and "Unnecessary". You will fill them in for each policy in Apply and Evaluate.
- Best after What Is Intelligence Management?.
- 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 in, on, out of the loop: Decides before the system acts; watches and can step in; not involved.

Escalation: Sending a case up to a more capable decision-maker.

Override and appeal: Changing a decision before it takes effect; challenging it afterwards.

Recourse: A real action in time, such as pause, request more evidence, or appeal.

Meaningful review: A review where the person could really change the outcome.

Automation bias: Going along with a machine's recommendation without really checking it.

Escalation policy: The rules that decide which cases a person must see.

Running the lesson 33 minutes

Open 4 min

"A reviewer has 8 seconds to click Confirm on an AI's denial. Did a person really decide? What would make that click count?"

Before anyone clicks, ask: is it safest to send every AI decision to a person? Take a show of hands, write the count on the board, and come back to it in Evaluate.

Step	What learners do	Ask the class	Listen for
1. Remember 5 min	Run Case 07, a housing denial, with nobody watching. Then seat a caseworker before, beside or after the decision, name each place (in, on or out of the loop) and sort four words.	• Seated after the decision, could the caseworker still change Case 07? • What is the difference between an override and an appeal?	Only a person before the outcome, or watching with power to stop it in time, can change it.
2. Understand 6 min	Split a review into five parts (information, time, context, authority, recourse), predict whether an 8-second confirmation click can change the outcome, then fix the desk until the review is real.	• The reviewer can click Confirm. Which of the five parts are missing? • Why was no single change enough to fix the desk?	A click is not control. The reviewer needs all five parts, and each missing part needs its own fix.
3. Apply 6 min	Learners switch on escalation rules (consequence, evidence gap, novel, rights, reversibility), run twelve simulated cases through the review queue, and open a case to see which rule sent it to a person.	• How many cases did your rules send to a person, and how many were missed? • Which rule would you add first, and what would it cost the reviewer?	Every rule has a cost: more cases caught means a longer queue for one reviewer.
4. Analyze 5 min	Learners diagnose two of three failures, each with marked evidence: a missed case under the Autopilot policy, a late review under Cautious, and a rubber-stamped review.	• Case 04 looked routine. Why did Autopilot let it through? • The same reviewer spent 92 seconds, then 4 seconds, on case 09. What changed?	A policy catches only what its rules test, and a long queue turns real review into rubber-stamping.
5. Evaluate 7 min	Learners compare two rival policies: Autopilot misses 3 cases; Cautious misses none but is late on 3. Then they build one with 0 missed, 0 late, at most 1 unnecessary, and defend its trade-off.	• Cautious missed nothing. Why is it still not safe? • Look back at our show of hands. Was sending everything to a person safest? • Which trade-off did your policy really make?	A policy defended with its own counts: life-safety cases first, or a narrow net so reviewers can really read.

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

Deepen the thinking

Critical thinking	A log says "reviewed by a person" for both of case 09's reviews. What else would you need to know before trusting that line?
AI thinking	Case 03 looked alarming but needed nobody; case 04 looked routine and hurt a family. What should an escalation rule test, if not how a case looks?
Ethics and fairness	An appeal that ends after support has already stopped is a right on paper. Who is harmed most when appeals come too late?
Connect to life	At school, who can change a decision once it is made, such as a grade or a team pick? Is that an override or an appeal?

Misconceptions to listen for

"A human clicked approve, so a human decided."

Understand's confirmation click has no evidence, an 8-second countdown and only a Confirm button. The click is real; the oversight is not.

"Sending everything to a person is the safe answer."

Cautious misses nothing, yet it is late on all three life-safety cases, because they wait behind routine work in the queue.

"If it is flagged, a person will catch it."

The replay shows one reviewer on case 09: 92 seconds with the evidence open, then 4 seconds unopened under a heavy queue.

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

- A referee watching a video replay has real oversight only with clear footage, time and power to change the call.
- Online platforms mix AI flags with human reviewers, and a rushed reviewer may just agree with the AI.
- In many places, you can ask for a person to review an automated decision about you: an appeal.

Try it this week. Next time you click "OK" without reading, ask: did I have the information, time and power to really decide?

Watch out for. A person clicking "approve" isn't real oversight without information, time and power. And sending everything to people makes urgent cases wait.

Talk about it. Is it always safer to have a person check everything? What could go wrong?

Exit ticket

1. Name the five parts a reviewer needs for a review to be meaningful.

   _____

2. What is the difference between human in the loop and human on the loop?

   _____

3. Why can sending more cases to people make some cases worse?

   _____

Going further

In the Open Project Canvas Lab, learners add an "Ask a person" state to the Agent Loop that moves on only when they click the stage. Is that human in or on the loop?