STIAEM

Slow Tech Innovation

AI, Ethics and

New Teaching Methods

Operational protocols for the 'Human-in-the-Loop' model in educational governance

Co-funded by the European Union

2025-1-IT02-KA220-SCH-000365213

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Education and Culture Executive Agency (EACEA). Neither the European Union nor EACEA can be held responsible for them... Project reference code: 2025-1-IT02-KA220-SCH-000365213

Università di Foggia

Studi Umanistici, Lettere, Beni Culturali, Scienze della Formazione

İSTANBUL - ESENLER İLÇE MİLLİ EĞİTİM MÜDÜRLÜĞÜ

1474

STIAEM
Slow Tech Innovation
AI, Ethics and
New Teaching Methods

Co-funded by
the European Union

2025-1-IT02-KA220-SCH-000365213

eduOpen



CC BY-NC-SA 4.0 International

This license requires that reusers give credit to the creator. It allows reusers to distribute, remix, adapt, and build upon the material in any medium or format, for noncommercial purposes only. If others modify or adapt the material, they must license the modified material under identical terms.

Università di Foggia
Studi Umanistici, Lettere, Beni Culturali,
Scienze della Formazione

OER 3.2.2: Operational protocols for the 'Human-in-the-Loop' model in educational governance
Reference Module: Module 3 - AI Ethics and Systemic Biases
Related Video-Pill: Video 3.2 - Accountability and Human-in-the-Loop
Authoring partner: Esenler District Directorate of National Education - Türkiye
Topic / Thematic Focus: Operational protocols, 'Human-in-the-Loop' model and educational governance

Introduction and Learning Objectives

This resource completes, on the operational level, the path begun with OER 3.2.1 and accompanies video-pill 3.2 "Accountability and Human-in-the-Loop". Whereas there it was clarified who is accountable and why, here the how is addressed: which protocols to adopt so that human oversight does not remain a statement of principle but becomes a verifiable practice.

The reference framework is the Human-in-the-Loop (HITL) model as described by Mosqueira-Rey and colleagues (2023), read in the light of the European regulation on AI and of the OECD and UNESCO principles. The connection with DigComp 3.0 (Cosgrove & Cachia, 2025) and with the UNESCO framework for teachers (Miao & Cukurova, 2024) places the protocol within professional competence, not outside it.

By the end of the reading, teachers will be able to:

distinguish HITL supervision models (active learning, interactive machine learning, machine teaching) according to who controls the process;

apply a three-phase protocol - before adoption, during use, after evaluation - with defined actions and responsibilities;

identify the conditions of durability (training, traceability, oversight inside the cycle) that make the protocol sustainable.

1. Not all oversight is the same: who has control?

The term "human-in-the-loop" is often used generically. Research shows, on the contrary, that oversight takes different forms, and these differences matter. Mosqueira-Rey and colleagues distinguish the types of human-machine interaction according to who controls the learning process:

"Depending on who is in control of the learning process, we can identify: active learning, in which the system remains in control; interactive machine learning, in which there is a closer interaction between users and learning systems; and machine teaching, where human domain experts have control over the learning process" (Mosqueira-Rey et al., 2023).

In active learning the system remains in control and calls on the human as an "oracle" to label uncertain cases. In interactive machine learning there is a denser exchange: the user supplies information in a more focused and incremental way. In machine teaching the human expert controls the process: selecting the examples, setting the criteria, deciding the sequence. The guiding question for schools is therefore: in which position do we want to be? The model that best protects professional judgement is the one in which it is the teacher who holds control.

"Active learning (AL), in which the system remains in control of the learning process and treats humans as oracles to annotate unlabeled data" (Mosqueira-Rey et al., 2023).

"Interactive machine learning (IML), in which there is a closer interaction between users and learning systems, with people interactively supplying information in a more focused,

frequent, and incremental way compared to traditional machine learning" (Mosqueira-Rey et al., 2023).
"Machine teaching (MT), where human domain experts have control over the learning process by delimiting the knowledge that they intend to transfer to the machine learning model" (Mosqueira-Rey et al., 2023).
For schools the difference is not terminological. In active learning the teacher is reduced to a supplier of labels, and the direction remains with the machine; in interactive machine learning the teacher co-designs and corrects; in machine teaching it is the teacher who defines what the machine must learn. The choice among these models is not technical but pedagogical: it tells us who holds control of the educational process.

2. Before adoption: read, assess, document

The protocol begins before the tool enters the classroom. The first action is to read the technical documentation: which data were used for training, which validity metrics are declared, which limitations are admitted. The second is the impact assessment: verifying how the system behaves with respect to one's own school population, and in particular whether distortions exist for specific groups. The third is the documentation of decisions: who adopts, with what pedagogical rationale, which alternatives were considered.
This moment corresponds, in the European framework, to the deployer's obligations: it is a matter of ensuring that input data are relevant and representative and that the system is used in accordance with the instructions. This is not bureaucracy: it is the condition for oversight subsequently to be exercised with full awareness.

3. During use: exercising the safeguards and keeping the register

In the midst of use, oversight takes concrete form in the exercise of the powers provided for by Article 14 and in their recording. The teacher understands the system's capabilities and limitations, monitors anomalies and unexpected performance, interprets the output taking into account the context the machine does not see, and has the possibility of disregarding, altering or interrupting the recommendation.

Alongside the safeguards, the protocol requires a memory. The pedagogical register notes, for each assisted decision: the moment, the tool, the output produced, the human decision and the rationale. It also notes anomalies, false alarms and correction errors. This memory has a double function: it makes what happens traceable - an accountability requirement - and it feeds organisational learning, because recurring error patterns become training material.

"Deployers shall entrust human oversight to natural persons who have the necessary competence, training and authority, as well as the necessary support" (Reg. EU 2024/1689, Art. 26(2)).

"Deployers shall monitor the operation of the high-risk AI system on the basis of the instructions for use and, where relevant, inform providers thereof..." (Reg. EU 2024/1689, Art. 26(5)).

These obligations have an immediate organisational translation. "Competence, training and authority" are not abstract requirements: they indicate that the school must appoint those responsible for oversight, train them adequately and confer on them the authority to intervene. Oversight entrusted to someone with no say in the decisions is, in practice, no oversight at all.

4. After evaluation: collegial review and aggregated data

The third phase turns individual experience into collective decision. The review should not be conducted case by case on individual students, but on aggregated data: how many outputs were accepted, how many modified, how many rejected; in which subjects; with which types of error. From this follow the institution's policy guidelines: which tools to confirm, which to suspend, which precautions to strengthen.

This moment realises the most mature part of accountability: rendering account not only to an external authority, but to the educational community. And it is also the place where it is decided whether and how to inform students and families, in coherence with the UNESCO (2021) principles of transparency and participation.

"AI actors should implement mechanisms and safeguards, such as capacity for human agency and oversight, including to address risks arising from uses outside of intended purpose, intentional misuse, or unintentional misuse in a manner appropriate to the context and consistent with the state of the art" (OECD, 2024, principle 1.2, letter b).

"Mechanisms should be in place, as appropriate, to ensure that if AI systems risk causing undue harm or exhibit undesired behaviour, they can be overridden, repaired, and/or decommissioned safely as needed" (OECD, 2024, principle 1.4, letter b).

The principle of human agency links oversight to the concrete possibility of intervening: it is not enough to observe, one must be able to override, repair, decommission. In the school protocol this means providing, already at the adoption stage, for the ways in which the teacher can suspend or cancel an automatic outcome, and who is accountable for that decision.

The UNESCO recommendation on the ethics of AI recalls the duty to govern the human impacts of systems (UNESCO, 2021).

5. The conditions of durability: training, traceability, oversight inside the cycle

A protocol is useful only if it holds over time. Three conditions guarantee its durability. The first is continuous training: European obligations require that oversight be entrusted to personnel with adequate competence and authority, and competence is built, not presupposed. The second is traceability: without logs and documentation, oversight becomes a memory, not a proof. The third, the most important on the pedagogical plane, is oversight inside the cycle: oversight cannot be reduced to a final check, it must operate while the process unfolds.

The OECD (2024) principle that systems must be able to be overridden, repaired or decommissioned safely, and that traceability of datasets, processes and decisions is required, also applies to schools, although the school is not a technology provider. What makes the difference is the positioning of the teacher: inside the cycle, not downstream. This is the formula to keep in mind: oversight is inside the cycle, not afterwards.

A final warning concerns the risk of bureaucratic drift: a protocol rich in forms but lacking pedagogical substance protects no one. The measure of its effectiveness is not the number of documents produced, but the school's ability to recognise and correct a system error before it becomes harm to a student. For this reason oversight must be kept anchored to the concrete case, without losing sight of the person the decision concerns.

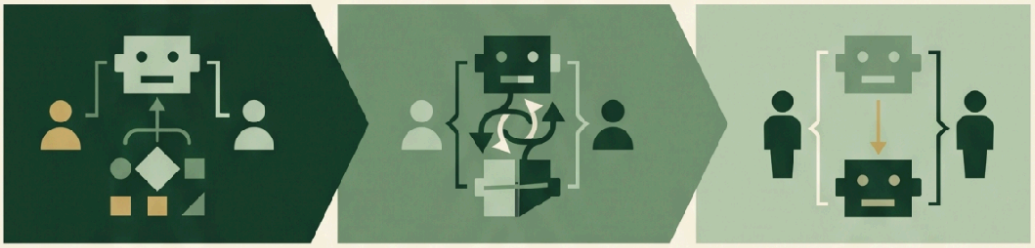
Phase	Key actions	Responsible	Tools
Before adoption	Reading documentation, assessing impact, documenting the choice	Institution + lead teacher	Guidelines, data analysis, minutes
During use	Exercising the safeguards, interpreting, recording anomalies	Teacher	Pedagogical register, system logs
After evaluation	Reviewing aggregated data, updating the policy guidelines	School board / team	Aggregated reports, policy guidelines

Visual Scaffolding

OPERATIONALIZING HUMAN-IN-THE-LOOP: A Protocol for Educational Governance

A verifiable 3-phase operational protocol for maintaining human oversight and accountability when using AI systems in schools.

TAXONOMY OF CONTROL: WHO IS IN THE LOOP?



ACTIVE LEARNING
(System in Control)
The AI system drives the process, treating the teacher as an "oracle" only to label or annotate data the machine finds uncertain.

INTERACTIVE MACHINE LEARNING
(Co-Design)
A denser exchange where the teacher supplies frequent, incremental information to correct and refine the system.

MACHINE TEACHING
(Human in Control)
The teacher acts as the domain expert, selecting examples, setting criteria, and delimiting the knowledge the machine is allowed to learn.

THE 3-PHASE OPERATIONAL PROTOCOL

A verifiable practice for educational governance and professional competence.



PHASE 1: BEFORE ADOPTION
(The "Deployer" Obligations)

- Schools must read technical documentation
- Assess impact on their specific student population to find biases
- Document the pedagogical rationale for choosing the tool.

PHASE 2: DURING USE
(Exercising Art. 14 Safeguards)

- Teachers monitor performance
- Interpret outputs in context
- Exercise the authority to disregard, alter, or interrupt AI recommendations.

PHASE 3: AFTER EVALUATION
(Collective Accountability)

- The school board reviews aggregated data—not individual cases
- To update policy guidelines
- Decide which tools to suspend, and inform families.

THE PEDAGOGICAL REGISTER & DECISION MEMORY

Oversight requires a memory to ensure accountability and organizational learning.



DOCUMENTING ASSISTED DECISIONS

For every decision made with AI, the register must record the tool used, the machine's output, the final human decision, and the rationale.



TRACKING ANOMALIES

Recording false alarms and system errors turns recurring patterns into training material for the entire school community.

THE CONDITIONS OF DURABILITY



CONTINUOUS TRAINING
Competence and authority are built, not presupposed; oversight must be entrusted to personnel with adequate pedagogical training.



TRACEABILITY
Without logs and evidentiary documentation, oversight is merely a statement of intent rather than a verifiable proof.



OVERSIGHT INSIDE THE CYCLE
Oversight must operate while the process unfolds (not post-hoc), ensuring the teacher is positioned "inside the cycle" of decision-making.

Conclusions and Metacognitive Reflection Prompts

The Human-in-the-Loop model becomes school governance only if translated into a protocol. Distinguishing who controls the process, defining actions for the three phases and ensuring the conditions of durability enables the school to use AI without delegating judgement to it. The pedagogical register and collegial review turn oversight from intention into verifiable practice, and traceability makes the institution able to render account of its own decisions.

To design your own protocol:

- In which of the three HITL models does your use of AI - actual or imagined - fall, and which would you prefer in order to protect your professional judgement?
- What should a pedagogical register of assisted decisions in your subject concretely contain?
- Which of the three conditions of durability (training, traceability, oversight inside the cycle) represents your weakest point today, and how could you strengthen it?

Technical Metadata Sheet

Metadata Field	Specification / Parameter Value
Identifier	OER 3.2.2
Title	Operational Protocols for the 'Human-in-the-Loop' Model in Educational Governance
Creator	Esenler District Directorate of National Education - Türkiye
Description	Complementing OER 3.2.1, the resource tackles the how of human oversight so that it becomes a verifiable practice rather than a statement of principle. It distinguishes HITL models (active learning, interactive machine learning, machine teaching) by who controls the process, and proposes a three-phase protocol — before adoption, during use, after evaluation. The practical focus is on the conditions of durability: training, traceability and oversight inside the cycle.
Keywords	human-in-the-loop; active learning; machine teaching; operational protocol; traceability; educational governance; oversight; institutional procedures
Language	EN (English)
Affiliation	Esenler District Directorate of National Education - Türkiye
Contributor	STIAEM Transnational Scientific Board (peer review, Rhodes validation protocol)
Media Format	Text; Infographics
Genre	Training Manual; Educational Guide
URL	
Citation	European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj Cosgrove, J., & Cachia, R. (2025). DigComp 3.0: European digital competence framework (Fifth edition). Publications Office of the European Union. https://data.europa.eu/doi/10.2760/0001149 Mosqueira-Rey, E., Hernandez-Pereira, E., Alonso-Rios, D., Bobes-Bascaran, J., & Fernandez-Leal, A. (2023). Human-in-the-loop machine learning: A state of the art. Artificial Intelligence Review, 56(4), 3005-3054. https://doi.org/10.1007/s10462-022-10246-w OECD. (2024). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). OECD. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO. https://doi.org/10.54675/ZJTE2084
Educational Level	Secondary School

Associated Module / MOOC Video Lesson	Module 3 - AI Ethics and Systemic Biases / Video 3.2 - Accountability and Human-in-the-Loop
Learning Objectives	distinguish HITL supervision models (active learning, interactive machine learning, machine teaching) according to who controls the process; apply a three-phase protocol - before adoption, during use, after evaluation - with defined actions and responsibilities; identify the conditions of durability (training, traceability, oversight inside the cycle) that make the protocol sustainable.
Rights	Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)
Date Available	14/09/2026
Date Modified	

References and Webography

European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union. https://eur-lex.europa.eu/eli/reg/2024/1689/oj

Cosgrove, J., & Cachia, R. (2025). DigComp 3.0: European digital competence framework (Fifth edition). Publications Office of the European Union. https://data.europa.eu/doi/10.2760/0001149

Mosqueira-Rey, E., Hernandez-Pereira, E., Alonso-Rios, D., Bobes-Bascaran, J., & Fernandez-Leal, A. (2023). Human-in-the-loop machine learning: A state of the art. Artificial Intelligence Review, 56(4), 3005-3054. https://doi.org/10.1007/s10462-022-10246-w

OECD. (2024). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). OECD. https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO. https://doi.org/10.54675/ZJTE2084

UNESCO. (2021). Recommendation on the ethics of artificial intelligence. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137