

Explainable AI in Education

Category: Artificial Intelligence in Education

Explainable AI in education refers to AI systems and practices that provide people with useful accounts of how outputs or recommendations were produced, what evidence influenced them and where uncertainty or limitations remain. An explanation should support understanding and action, not merely expose technical detail.

Different users need different explanations. A student may need to know why a resource was recommended; an educator may need the influential factors and confidence; a developer may require model-level diagnostics. Explanation quality, therefore, depends on purpose, audience, and consequence ([Khosravi et al., 2022](#)).

Explanation Requirements

1. **Purpose:** State what the system is designed to do and what the output may legitimately support.
2. **Influence:** Identify the principal data or factors that shaped a particular result.
3. **Uncertainty:** Communicate confidence, known limitations and situations in which performance may weaken.
4. **Actionability:** Help the user decide what to check, change or do next without manipulating behaviour.
5. **Contestability:** Provide a route to question the output, verify the data, and request human review.

Examples in Practice

A writing-support system highlights a paragraph as unclear. A useful explanation identifies the sentence features that influenced the judgement, shows the relevant criterion and offers alternatives. The student can reject the suggestion to preserve authorial intention and report a recurring misinterpretation.

Why It Matters

Explanations can support trust calibration, learning and accountable review, especially when AI informs feedback or access to support. Yet a plausible explanation does not prove that a model is accurate or fair; institutions must test both the system and people's understanding of its explanations ([European Commission, 2022](#)).

Further Reading

6. Khosravi, H., et al. (2022). [Explainable artificial intelligence in education](#). Computers and Education: Artificial Intelligence, 3, 100074.
7. European Commission. (2022). [Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators](#). Publications Office of the European Union.
8. Miao, F., & Cukurova, M. (2024). [AI competency framework for teachers](#). UNESCO.

Related Terms: AI Transparency | Human Oversight of AI | Algorithmic Literacy | AI Bias, Fairness and Inclusion