STIAEM — Open Educational Resource (OER) Repository for AI

Metadata Sheet — WP2 / Erasmus+ KA220-SCH, ref. 2025-1-IT02-KA220-SCH-000365213

OER 6.2.1

Dismantling False Controls: Limitations and False-Positive Rates of Anti-Plagiarism Software

Identifier	OER 6.2.1
Title	Dismantling False Controls: Limitations and False-Positive Rates of Anti-Plagiarism Software
Creator	Esenler District Directorate of National Education - Türkiye
Description	The resource examines a control that promises certainty and introduces error: the automatic detection of AI-generated text. It explains why detectors estimate statistical parameters and do not recognise AI, and describes, prudently and with attribution, the problem of false positives against non-native-speaker authors (Liang et al., 2023). The practical focus is on placing detectors in an exploratory function and excluding them from high-stakes decisions.
Keywords	AI-text detection; false positives; perplexity; burstiness; non-native speakers; measurement error; high-stakes decisions; academic integrity
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	https://www.stiaemproject.eu/?oer_download=OER-6.2.1
Citation	Esenler District Directorate of National Education. (2026). *Dismantling False Controls: Limitations and False-Positive Rates of Anti-Plagiarism Software* [Educational guide]. Slow Tech Innovation: AI, Ethics, and New Teaching Methods (STIAEM), 2025-1-IT02-KA220-SCH-000365213.
References	Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228-239. https://doi.org/10.1080/14703297.2023.2190148 Eaton, S. E. (2023). Postplagiarism: The future of academic integrity in the age of artificial intelligence [Academic debate on post-plagiarism]. University of Calgary. Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). GPT detectors are biased against non-native English writers. Patterns, 4(7), 100779. https://doi.org/10.1016/j.patter.2023.100779 Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. Journal of University Teaching & Learning Practice, 20(2). https://ro.uow.edu.au/jutlp/vol20/iss2/07/ UNESCO. (2023). Guidance for generative AI in education and research. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693
Educational Level	Secondary School
Associated Module /	Module 6 - Classroom Practice and Authentic Assessment / Video

MOOC Video Lesson	6.2 - Redefining Plagiarism in School
Learning Objectives	Explain why automatic detectors estimate statistical parameters and do not recognise AI; describe, prudently and with attribution, the problem of false positives against non-native-speaker authors; place detectors in their exploratory function, excluding them from high-stakes decisions.
Rights	Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)
Date Available	14/09/2026
Date Modified	