STIAEM — Open Educational Resource (OER) Repository for AI

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

OER 2.2.1

The Mechanics of Large Language Models: Deconstructing the 'Statistical Mirror'

Identifier	OER 2.2.1
Title	The Mechanics of Large Language Models: Deconstructing the 'Statistical Mirror'
Creator	University of Alicante - Spain
Description	The text makes comprehensible how a large language model generates text, deconstructing the metaphor of the statistical mirror. It explains autoregressive generation, few-shot learning and emergent capabilities, distinguishing formal coherence from factual accuracy. The didactic focus is on understanding why the model simulates rather than understands, and why scale produces statistical jumps and not mental faculties.
Keywords	large language models; autoregressive generation; statistical mirror; few-shot learning; emergent capabilities; factual accuracy; AI literacy; secondary school
Language	EN (English)
Affiliation	University of Alicante - Spain
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-2.2.1
Citation	University of Alicante. (2026). *The Mechanics of Large Language Models: Deconstructing the 'Statistical Mirror'* [Educational guide]. Slow Tech Innovation: AI, Ethics, and New Teaching Methods (STIAEM), 2025-1-IT02-KA220-SCH-000365213.
References	Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610-623). ACM. https://doi.org/10.1145/3442188.3445922 Bender, E. M., & Koller, A. (2020). Climbing towards NLU: On meaning, form, and understanding in the age of data. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 5185-5198). Association for Computational Linguistics. https://aclanthology.org/2020.acl-main.463/ Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877-1901. https://arxiv.org/abs/2005.14165 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 Miao, F., & Cukurova, M. (2024). AI competency framework for teachers. UNESCO. https://doi.org/10.54675/ZJTE2084 Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S.,

	... Fedus, W. (2022). Emergent abilities of large language models. Transactions on Machine Learning Research. https://arxiv.org/abs/2206.07682
Educational Level	Secondary School
Associated Module / MOOC Video Lesson	Module 2 - AI Fundamentals and LLM Mechanics / Video 2.2 - Probabilistic Functioning of Large Language Models
Learning Objectives	Describe the autoregressive generation of text and explain why the model chooses each word on the basis of probability; distinguish formal coherence from factual accuracy, recognising in the 'statistical mirror' a simulation and not an understanding; recognise, on the basis of few-shot learning and emergent capabilities, that scale produces statistical jumps and not mental faculties.
Rights	Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0)
Date Available	14/09/2026
Date Modified	