Word	Definition	Source
Ethical	**Ethical: Seeking accountability regarding fundamental human values and rights.** Ethical AI considers the rights and values of the people who are working with the AI or impacted by the AI, particularly within sensitive contexts.	Capel, T., & Brereton, M. (2023, April). What is human-centered about human-centered AI? A map of the research landscape. In Proceedings of the 2023 CHI conference on human factors in computing systems (pp. 1-23).
Accurate	**Accurate: An AI system which correctly performs its intended task i.e. produces correct outputs based on its input data.** This colloquial sense of "accuracy" is not limited strictly to the proportion of correctly classified instances (the technical definition), but can also encompass other metrics relevant to measuring desired model performance on the specific task, such as precision, recall, and F1-score.	Key Terms for AI Governance, iapp, July 2024, https://iapp.org/resources/article/key-terms-for-ai-governance/
Unbiased	**Unbiased: AI systems and models which do not display "systematic errors or unfair tendencies that lead to discriminatory outcomes"** The range of discriminatory outcomes that biased models can cause include both under- and over-representation in particular contexts, increased salience of particular stereotypes or harmful representations, and also capability biases (whereby the model performs more poorly on tasks related to certain demographic groups).	Ranjan, R., Gupta, S., & Singh, S. N. (2024). A Comprehensive Survey of Bias in LLMs: Current Landscape and Future Directions. arXiv preprint arXiv:2409.16430.
Trustworthy	**Trustworthy** can be thought of as a combination of being robust and being truthful Robustness refers to the ability of an AI system to maintain its performance and accuracy under different conditions, including in the face of adversarial attacks Truthfulness is the degree to which the AI system provides truthful and honest outputs	Page 5, Yang, Ge, et al. "LLMCBench: Benchmarking Large Language Model Compression for Efficient Deployment." arXiv preprint arXiv:2410.21352 (2024) https://arxiv.org/abs/2410.21352
Transparent	Transparent AI can refer to: • Developing and using AI systems in a way that allows "appropriate" traceability and explainability • Making humans aware that they are communicating or interacting with an AI system • Informing AI-deployers of the capabilities and limitations of that AI system • Informing persons by the AI system about their rights.	Artificial Intelligence Act, OJ L, 2024/1689, Recital 27