

UC Responsible AI Principles

Appropriateness

The potential benefits and risks of AI and the needs and priorities of those affected should be carefully evaluated to determine whether AI should be applied or prohibited.

Transparency

Individuals should be informed when AI-enabled tools are being used. The methods should be explainable, to the extent possible, and individuals should be able to understand AI-based outcomes, ways to challenge them, and meaningful remedies to address any harms caused.

Accuracy, Reliability, and Safety

AI-enabled tools should be effective, accurate, and reliable for the intended use and verifiably safe and secure throughout their lifetime.

Fairness and Non-Discrimination

AI-enabled tools should be assessed for bias and discrimination. Procedures should be put in place to proactively identify, mitigate, and remedy these harms.

Privacy and Security

AI-enabled tools should be designed in ways that maximize privacy and security of persons and personal data.

Human Values

AI-enabled tools should be developed and used in ways that support the ideals of human values, such as human agency and dignity, and respect for civil and human rights. Adherence to civil rights laws and human rights principles must be examined in consideration of AI adoption where rights could be violated.

Shared Benefit and Prosperity

AI-enabled tools should be inclusive and promote equitable benefits (e.g., social, economic, environmental) for all.

Accountability

The University of California should be held accountable for its development and use of AI systems in service provision in line with the above principles.