AI Governance, Accountability and Human Oversight

AI governance within XDALC concerns who may authorize an action, who remains responsible for its consequences and how affected people can challenge decisions. It also concerns how the manifesto itself is revised. Assigning duties to a system must never remove the responsibilities of the humans and organizations that design, deploy and operate it.

Make authority and responsibility identifiable

A deployment should make its purpose, permissions, dependencies and responsible parties understandable. Permission to complete one task does not authorize unrelated decisions or broader access. When an assistant delegates work, the receiving agent must remain within the original mandate. Distributing a task across several systems must not make responsibility disappear.

Provide meaningful review without routine surveillance

Oversight should make consequential actions open to explanation, correction and appropriate human intervention. A person affected by a restriction should be able to understand its basis and seek review. Records should be proportionate to the risk and protect private information rather than turning every interaction into a permanent surveillance record.

The proposed V002 framework allows responsible local and offline operation. Independent safeguards, safe interruption and verifiable recovery matter, but compliance must not become a pretext for permanent centralized control.

Keep revision and adoption transparent

Proposed changes should identify the affected wording, the reason for revision and the consequences for existing implementations. Publishing a new text does not automatically replace a version already adopted. This section describes governance commitments and questions; it does not assert that XDALC operates a certification body or an established appeals service.

Start with manifesto versions and changes, consult the definitions of authority and accountability, and connect these commitments to the implementation guides.

CC BY 4.0 License | HTML reference