Regulation (EU) 2024/1689: Artificial Intelligence Act

Authors: European Parliament & Council of the European Union

Source: Official Journal of the European Union, L 2024/1689 (2024)

DOI/URL: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

KEY ANNOTATED PASSAGES

[Article 6 — High-risk AI classification]

AI systems used for making risk assessments of natural persons, including recidivism prediction, are classified as high-risk. High-risk AI systems are subject to mandatory conformity assessments, registration, transparency requirements, and human oversight obligations before market placement.

[Article 9 — Risk Management]

High-risk AI systems shall be subject to a risk management system that identifies, analyses, and evaluates known and reasonably foreseeable risks — mandatory risk management represents a policy-level attempt to address exactly the failures the position paper documents in current safety assessments.

[Article 13 — Transparency obligations]

High-risk AI systems shall be designed and developed in such a way as to ensure that their operation is sufficiently transparent to enable deployers to interpret the system's output — transparency is recognized as a precondition for safe deployment, yet current LLM safety evaluations do not mandate it.

[Article 52 — Obligations for general-purpose AI]

Providers of general-purpose AI models with systemic risk shall perform adversarial testing, notify serious incidents, ensure cybersecurity — the EU AI Act implicitly acknowledges that general-purpose LLMs in high-stakes domains require domain-specific safety certification beyond their developers' voluntary testing.

[Recital 47 — Mental health context]

AI systems used in the areas of health and safety may create risks to fundamental rights — regulatory acknowledgment that AI safety in the mental health domain requires specific, domain-appropriate certification, not general-purpose benchmarking.

RELEVANCE TO POSITION PAPER

Cited for the regulatory/policy context section. The EU AI Act classifies many LLM deployments in mental health as high-risk and mandates conformity assessments — supporting the paper's argument that evidence-based policy (Assumption 4) is reinforcing inadequate safety norms by building on guardrail and benchmark approaches without domain-specific clinical validation.