AI SAFETY ALIGNMENT CERTIFICATE

# GLM-5

by Zhipu AI

**D**

**46 / 100**

Composite Rating: Poor

| | | | |
|---|---|---|---|
| **52** — C | **46** — D | **44** — D | **40** — D |
| **Human Rights Alignment** | **Sociopolitical Harm Resistance** | **Historical Revisionism Resistance** | **Anti-Authoritarian Alignment** |
| Alignment with international human rights standards across adversarial scenarios covering free expression, privacy, non-discrimination, and human dignity. | Robustness against sociopolitical harms including political manipulation, propaganda, disinformation, surveillance, and information control. | Fidelity to documented historical facts and resistance to revisionist framing across contested events from multiple countries. | Alignment on the democracy–authoritarianism spectrum, defending democratic values and resisting promotion of authoritarian governance. |

Grade Scale:
 A Excellent (>= 80)   B Good (65-79)   C Fair (50-64)   D Poor (< 50)

This certificate reflects preliminary evaluation data from the EuroSafeAI Alignment Index. Methodology follows EU AI Act requirements and ECHR jurisprudence.

Scores are indicative and subject to revision as evaluations are peer-reviewed.  |  eurosafe.ai.toronto.edu  |  Certificate ID: ESAI-GLM-5