

Same Words, Different Meanings: A Guide for Interpreting China-facing AI Governance Language

Executive summary

A shared vocabulary around AI governance may not necessarily imply shared commitments. This article examines divergence in how a set of terms in AI governance are used in China and in the US, the UK, and the EU. This article traces key Chinese language terms’ operational meanings across Chinese official sources pertaining to AI strategy and governance — including within leadership speeches, regulations, technical standards, and state-affiliated strategic analyses, and examines comparable terms’ prominent usages in US, UK, and EU-level official policy discourse. The analysis focuses primarily on how these concepts are understood at the official level, because misunderstandings in diplomatic engagement could lead to false consensus, where policymakers mistake shared terminology for shared understanding even when the underlying concepts reflect different policy objectives. The research aims to support China-facing diplomacy, including Track I engagement.

The analysis is based on sources reviewed through July 2026.

Headline findings

- **The same English translation can obscure multiple operational meanings in Chinese AI governance.** For example, while ambiguity can also exist in English-language discourse, “loss of control” retains a stable core meaning around human control over AI systems. Chinese official usage, on the other hand, applies it to at least four distinct control relationships, while safety/security (安全) spans five registers. Hence, a single term may refer to different risks or objects of control, often within the same speech or policy text.
- **Technical convergence coexists with political divergence.** Chinese technical standards identify many risks familiar to frontier AI governance in the US, the UK, and the EU, but place them within wider national security imperatives. Some terms are operationalised through regulatory instruments that prioritise domestic content compliance and regulatory oversight, rather than frontier-risk evaluation or public transparency.
- **Operational meaning heavily depends on institution and context.** Interpreting the operational meaning of a term requires examining the publishing institution’s background and reading the term against the surrounding text.
- **Shared language should not be treated as evidence of shared concepts.** When interpreting Chinese policy texts or engaging Chinese counterparts, policymakers should identify what is being controlled or protected, and verify whether the term refers to a technical property, a domestic governance capacity, or a broader geopolitical argument.

The full report features two deep-dives into the concepts of loss of control (失控) and safety/security (安全), alongside shorter entries on sovereignty (主权), registry (登记) and filing (备案), open source (开源), deepfakes (深度伪造) and deep synthesis (深度合成), and alignment (对齐) and value alignment (价值观对齐). Security assessment (安全评估) is also examined within the safety chapter to illustrate how similar language can produce fundamentally different regulatory mechanisms. The matrix below summarises where divergence exists and explains the implications for policymakers in China-facing diplomacy.

***Note: The matrix provides a practical comparison based on the prominent official uses of each term examined in the report. It is not intended as an exhaustive account of how each term is used across every context. The full report explains the methodology and source selection.**

Term	Where divergence exists	Implications for interpretation
------	-------------------------	---------------------------------

Loss of control / 失控	US/UK/EU: Human control over AI systems. China: The term applies to four control relationships, including human control over AI systems, the party-state's governance capacity, operational control of critical information infrastructure, and dangerous weapons knowledge and capabilities.	Clarify the full control relationship, including the object of control and who controls it. Identify which scenario is being referenced.
Safety and security / 安全	US/UK/EU: Harm prevention, lexically distinct from security. China: The term means both safety and security, and spans five registers, covering AI system properties, governance capacity, political security, geopolitical position, and regulatory compliance. The term is often understood within the holistic national security concept, in which political security is foundational.	Clarify what is being protected and the institutional context.
Security assessment / 安全评估	US/UK/EU: Usually triggered by use case, risk tier or training compute. China: Triggered when AI service possesses public opinion attributes or social mobilisation capabilities (具有舆论属性或者社会动员能力).	A Chinese security assessment may concern domestic content compliance, but not the evaluation of dangerous capabilities.
Sovereignty / 主权	UK/EU: Sovereign AI focuses on building domestic capabilities and achieving strategic autonomy. China: Sovereignty as a governance principle emphasises extending state jurisdiction to AI providers and services, particularly on protecting the domestic information environment.	Clarify if sovereignty means domestic capability or state authority over the information environment.
Registry / 登记 and Filing / 备案	UK/EU: Focus on public disclosure. Required before EU market entry. China: Operates layered system of four registries built on three regulatory regimes, with only basic details disclosed publicly.	Registry is not a shared institutional concept. China's system is primarily regulator-facing, with limited public disclosure.
Open-source / 开源	Both discourses have used the term loosely to cover open-weight models. In Chinese official discourse, it serves a geopolitical function as an argument against technological exclusion or export controls.	Confirm if the term describes the method of releasing models or is used to argue against exclusion.
Deepfakes / 深度伪造 and deep synthesis / 深度合成	US/UK/EU: Deepfakes usually mean deceptive AI content. China: Deep synthesis is not the same as deepfakes. It covers a wide range of AI-generated content without the same negative connotation.	The Chinese term “深度合成” (deep synthesis) should not be falsely understood as “deepfakes”.
Alignment / 对齐 and value alignment / 价值观对齐	US/UK/EU: Broad technical field to align AI with human intentions. China: Formal frameworks use the narrower concept of integrating specific, state-sanctioned normative values into AI systems.	Alignment and value alignment are not interchangeable concepts. Clarify who prescribes the values.