

Large Model and AIGC Safety Specification

Kuaishou Technology (hereinafter referred to as "Kuaishou" or "Company," together with its subsidiaries and consolidated affiliated entities, collectively referred to as "the Group"), in order to regulate the secure use of the Company's large models, effectively prevent security risks—including data leakage, resource misuse, data poisoning, and training data contamination—and to meaningfully safeguard the information security of the Company's large models, this Policy is hereby formulated in consideration of the Company's actual business operations.

This Policy applies to all employees of Kuaishou, as well as third-party personnel providing services to the Company (including, but not limited to, suppliers, outsourced personnel, interns, and others).

The Company has established an **Information Security Committee** as its highest decision-making body. The Committee is responsible for coordinating resources across all departments to respond to security crises and for making determinations on material matters relating to large model security development. The **Information Security Center** is responsible for execution and implementation: it shall, in accordance with the Committee's requirements, formulate large model security specifications, specific strategies, and process standards, and shall ensure their effective implementation and continuous improvement. Concurrently, it shall conduct periodic large model security testing and auditing to comprehensively safeguard the compliance and stable operation of business activities.

Data Usage Principles

In the course of developing and utilizing artificial intelligence, the Company shall at all times adhere to the "**Minimum Necessary Principle**," processing only the minimum scope of data required to support large model operations or to meet partner requirements.

With respect to data usage:

- Internal data inputs shall be processed strictly in accordance with their designated purposes; processing beyond the authorized scope is strictly prohibited.
- Sensitive personal information and trade secrets must undergo **desensitization** prior to processing.
- At the output stage, **security filtering mechanisms** shall be deployed to prevent the exposure of sensitive content.

With respect to data storage:

- The Company shall rigorously implement data classification and categorization standards together with confidentiality-level guidelines.
 - Data confidentiality levels shall be accurately assigned.
 - Authenticated access controls shall be enabled for sensitive data.
 - Sensitive personal information shall be subject to **encrypted storage**.
-

Model Usage Boundaries

The Company shall clearly define the authorized scope of use for artificial intelligence, establish permissible-use policies, deploy guardrail measures and human escalation mechanisms, and continuously monitor usage to prevent applications extending beyond the predetermined scope.

Network Security Protection

The Company has established a robust network security defense framework.

- **During the development phase**, employees shall strictly adhere to front-end secure coding standards to prevent the introduction of code security vulnerabilities at the source.
 - **Prior to release**, any integration of external large model services, initial deployment, and subsequent functional changes must undergo and pass a pre-release security assessment before formal launch.
 - **When establishing partnerships with third-party large model service providers**, the security responsibilities and obligations concerning data protection must be clearly defined, and detailed descriptions of their security measures must be obtained to sufficiently verify their capability to safeguard the Company's data security.
 - **User grievance channels**: Systems shall provide accessible portals for user appeals, public complaints, and incident reporting, ensuring that relevant feedback is promptly received and efficiently addressed.
-

Training Data Requirements

When utilizing proprietary or third-party large models for model training or fine-tuning:

- The origin and provenance of training datasets must be lawful and compliant; the use of unauthorized data shall be avoided.
- To safeguard the diversity of training data, content containing illegal, biased, or discriminatory elements shall be filtered and removed at the source during the training process.

- Where training data involves cross-border transmission, strict compliance with the relevant laws and regulations of the data's country of origin must be ensured, thereby building a comprehensive security and compliance framework for AI systems.
-

Content Moderation

Systems shall conduct content review at both the data input and synthesized output stages; the provision of illegal content is prohibited.

- The Company shall adopt a **combined technological and human review approach** to strengthen dual-gate oversight of user inputs and outputs within deep synthesis services.
 - A database of illegal and undesirable information characteristics shall be established and maintained, together with well-defined standards, rules, and procedures for inclusion. Relevant network logs shall be rigorously recorded and retained to support manual auditing.
 - Intervention mechanisms shall be activated immediately when required, with prompt measures taken—including halting content generation, suspending transmission, and removing offending content—to ensure that systems remain at all times under effective human oversight and control.
-

Environmental Footprint of AI Data Centers and Models

The Company is committed to reducing the ecological footprint of its AI data centers and models.

- **Energy efficiency management:** A systematic incentive and accountability mechanism shall be established, with Power Usage Effectiveness (PUE) indicators incorporated into departmental and employee performance evaluations.
- **Self-built data centers** shall deploy green technologies—including advanced cooling systems, energy-efficient building design, innovative power supply solutions, and periodic maintenance—to continuously improve resource utilization rates and maintain a low carbon footprint. Concurrently, periodic environmental monitoring shall be conducted to ensure that indicators for emissions, wastewater, noise, and other parameters remain superior to national standards.
- **For leased data centers**, priority shall be given to operators certified under environmental management system standards such as ISO 14001. An incentive and penalty mechanism shall be implemented to encourage operators to pursue low-carbon operations and to comprehensively fulfill their environmental responsibilities.

The Company shall periodically review this Policy and shall revise it as and when necessary. This Policy was first publicly disclosed to stakeholders in September 2026.