

ICS 35.240
CCS L 70



**NATIONAL STANDARD OF THE
PEOPLE'S REPUBLIC OF CHINA**

GB/Z 201-2026

**Artificial intelligence - Large model selection and application
guideline**

人工智能 大模型选型和应用指南

Issued on: July 30, 2026

Implemented on: July 30, 2026

**Issued by: State Administration for Market Regulation;
Standardization Administration of China**

Table of Contents

Foreword	III
Introduction	IV
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviated terms	1
5 Principles and process of selection and application	2
5.1 Principles	2
5.2 Process	2
6 Preparation for selection	3
6.1 Strategic planning	3
6.2 Organizational provision	3
6.3 Provision of personnel	4
6.4 Assessment of the present state	4
6.5 Preliminary selection of the model	5
7 Preparation of data	5
7.1 Data collection	5
7.2 Data storage	5
7.3 Data processing	6
7.4 Building the data set	6
8 Selection and optimization	6
8.1 Model selection	6
8.2 Model training	7
8.3 Model tuning	7
9 Deployment and application	8
9.1 Choice of deployment mode	8
9.2 Model deployment	8
9.3 Model integration	8
9.4 Secure operation and maintenance	8
10 Application evaluation	9
10.1 Model evaluation	9
10.2 Benefit evaluation	9

11 Risk management	9
11.1 Building organizational capability	9
11.2 Objectives of risk management	9
11.3 Risk identification and assessment	9
11.4 Risk response	10
Annex A (informative) Reference information for the selection of a large model	11
A.1 Estimation of computing power	11
A.2 Classification of models	11
A.3 Model parameters	11
A.4 Model performance	12
A.5 Model compression	12
Bibliography	14

1 Scope

This document establishes the principles of selecting an artificial intelligence large model and the process of applying it, covering the principles and process of selection, the preparation for selection, the preparation of data, selection and optimization, deployment and application, application evaluation and risk management, so as to give users of every kind guidance over the whole life cycle.

This document applies to the selection and application of artificial intelligence large model technology.

2 Normative references

The following documents contain provisions which, through normative reference in this text, constitute indispensable provisions of this document. For dated references, only the edition cited applies. For undated references, the latest edition, including any amendments, applies.

GB/T 23011-2022, Integration of informatization and industrialization - Digital transformation - Value and benefit reference model

GB/T 35273-2020, Information security technology - Personal information security specification

GB/T 36073-2025, Data management capability maturity assessment model

GB/T 37964-2019, Information security technology - Guide for de-identification of personal information

GB/T 45081-2024, Artificial intelligence - Management system

GB/T 45288.1-2025, Artificial intelligence - Large models - Part 1: General requirements

GB/T 45288.2-2025, Artificial intelligence - Large models - Part 2: Evaluation indicators and methods

GB/T 45288.3-2025, Artificial intelligence - Large models - Part 3: Service capability maturity assessment

3 Terms and definitions

The terms and definitions given in GB/T 23011-2022, GB/T 36073-2025, GB/T 37964-2019, GB/T 45081-2024, GB/T 45288.1-2025, GB/T 45288.2-2025 and GB/T 45288.3-2025 apply to this document.

4 Abbreviated terms

The following abbreviated terms apply to this document. API: Application Programming Interface. CPU: Central Processing Unit. DCMM: Data Capability Maturity Model. FLOPS: Floating-point Operations Per Second. GPTQ: Generative Pre-trained Transformer Quantization. GPU: Graphics Processing Unit. IT: Information Technology. LoRA: Low-Rank Adaptation. MaaS: Model as a Service.

5 Principles and process of selection and application

The basic process of selecting and applying a large model is shown in Figure 1, which sets risk management alongside the successive stages of preparation for selection (strategic planning, organizational provision, provision of personnel, assessment of the present state and preliminary selection of the model), preparation of data (data collection, storage, processing and building the data set), selection and optimization (model selection, model training and model tuning), deployment and application (deployment mode, model deployment, model integration and secure operation and maintenance), and application evaluation (model evaluation and benefit evaluation).

The selection and application of a large model comprise the following stages. a) Preparation for selection: the user needs to strengthen the stages of strategic planning, organizational provision, provision of personnel, assessment of the present state and preliminary selection of the model; the basic conditions for introducing artificial intelligence at present are assessed through market research, exchanges with peers, case analysis and trial experience; and, taking cost and expected benefit together, the target scenario and the medium- and long-term plan for introducing artificial intelligence are determined.

b) Preparation of data: the user strengthens its own capability to collect, store and process data; where necessary an external sector-specific data set may be purchased, or

techniques such as data inversion and simulated synthesis may be used, so as to build a data set proper to the user. c) Selection and optimization: according to the specific application scenario, a suitable model is screened along several dimensions such as performance and cost, and its effectiveness in the particular scenario is raised by targeted optimization through prompt engineering, retrieval-augmented generation, fine-tuning or comprehensive tuning.

d) Deployment and application: the manner of deployment is determined according to the technical capability of the user, the scale of system deployment, the security needs of the business and the overall running cost, and sufficient human, financial and material provision is made for the secure and stable running of the business. e) Application evaluation: a comprehensive evaluation of the effectiveness and the benefit of the applied model is organized periodically, and the overall plan for applying artificial intelligence is adjusted in the light of the development trend of the sector. Note: the process of each stage in Figure 1 is tailored to the actual circumstances.

6 Preparation for selection

6.1 Strategic planning. The user should include the artificial intelligence strategy in its overall business planning, making clear the objectives, the key fields and the route of implementation of the artificial intelligence application; should assess periodically the effect of artificial intelligence technology on its existing business processes, organizational structure and business model; and should draw up the corresponding transformation strategy, so as to ensure that the artificial intelligence application is consistent with its own long-term development objectives.

6.2 Organizational provision. The user should establish an organizational structure suited to the artificial intelligence application and set up a dedicated artificial intelligence project team, including the roles of project manager, data scientist, algorithm engineer and business analyst, with the responsibilities and authority of each role made clear. A leading group headed by the principal or a responsible officer of the enterprise should be established.

7 Preparation of data

7.2 Data storage. a) Storage, efficient access and backup and recovery are provided so as to prevent the loss or corruption of data. b) Data is collected, transmitted and stored according to the requirements of Clauses 5 and 6 of GB/T 35273-2020, and effective technical measures are used to protect the personal information in the data collected.

7.3 Data processing. The user carries out data processing in the following respects. a) Data cleaning: data quality problems in the data collected, such as anomalies, redundancy and missing values, should be dealt with by cleaning and conversion, so as to raise the quality