ICS 35.160
CCS L 62

NATIONAL STANDARD OF THE PEOPLE'S REPUBLIC OF CHINA

GB/Z 177.2-2026

Intelligence grading of artificial intelligence terminal - Part 2: General requirements

人工智能终端智能化分级 第2部分：总体要求

Issued on: April 30, 2026 **Implemented on: October 1, 2026**

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

Table of Contents

Foreword

This document was drafted in accordance with the rules given in GB/T 1.1-2020, Directives for standardization - Part 1: Rules for the structure and drafting of standardizing documents.

This document is Part 2 of GB/Z 177, Grading of intelligence for artificial intelligence terminals.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. The issuing body of this document is not to be held responsible for identifying any or all such patent rights.

This document was proposed by and is under the jurisdiction of the National Information Technology Standardization Technical Committee (SAC/TC 28).

1 Scope

This document specifies the system for grading the intelligence of artificial intelligence terminals, referred to below as the terminal, covering the division into intelligence levels, the capability elements and the key capabilities of each level, and describes the test methods.

This document applies to the grading of intelligence of terminals of every kind, and provides a reference for the design, development, application, selection and testing of terminals.

2 Normative references

The following document contains 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 of the referenced document, including any amendments, applies.

GB/Z 177.1, Grading of intelligence for artificial intelligence terminals - Part 1: Reference framework

3 Terms and definitions

For the purposes of this document, the terms and definitions given in GB/Z 177.1 and the following apply.

3.1 testing scenario - a composite test situation set up, within a particular application scenario, to test the key capabilities of a terminal.

3.2 testing task - the basic unit of execution in the testing of terminal intelligence. Note: a testing task can be a specific step within a testing scenario, or can be completed on its own as an independent, atomic test item.

3.3 scenario context - the dynamic set of information about the user, the device and the environment that a terminal builds in real time while carrying out a particular task, so as to understand and respond to the intent of the user accurately.

3.4 end-to-end closed-loop - the complete flow from the moment the user issues an instruction or an intent, through the perception, cognition and execution the terminal must

perform, to the successful delivery of the result of the task.

3.5 personal knowledge base - the set of personalized information about a user - particular facts, relationships, preferences and habits - that a terminal learns and stores, with the user's authorization, over long-term use.

3.6 session - one process of interaction between the user and the terminal to complete a task or a group of related tasks.

3.7 session context - the set of dynamic information recorded and maintained within one session - the interaction history and the current state - so as to keep the interaction coherent and the task accurately carried out.

3.8 ability element - the basic categories that make up the top-level framework of terminal intelligence capability.

3.9 task decomposition - the capability of a terminal to break a complex intent or a multi-step instruction from the user automatically into a series of smaller, more specific and executable subtasks.

3.10 task orchestration - the capability, on the basis of task decomposition, to plan and generate an optimal or reasonable order of execution according to the logical relations, the dependencies and the priorities among the subtasks.

3.11 intent clarification - the capability, when the terminal recognizes that the intent of the user is unclear, to make that intent clearer by asking of its own accord, offering options or requesting further information.

4 Abbreviated terms

For the purposes of this document, the following abbreviated term applies. TTS: text-to-speech.

5 Overview

According to the complexity of the capability with which a terminal carries out tasks of various kinds within its functional range, and the degree of automation, the level of terminal intelligence is divided into four levels, L1 to L4. The higher the level, the higher the level of intelligence of the terminal. A terminal of a higher level covers all the capabilities of the terminals of lower levels: an L3 terminal, for instance, has at the same time all the key capabilities of L1 and of L2.

6 Division into levels

6.1 L1 response level. The terminal understands a single simple instruction, calls a determined tool according to that instruction and completes a single-step task.

6.2 L2 tool level. The terminal understands the instruction of the user and simple intents and has a simple reasoning capability; it can call preset tools and complete single-step or clearly defined multi-step execution tasks; it can generate content in at least one modality, text, audio or image; and it has short-term memory within a single session.

6.3 L3 assistance level. The terminal can understand the instruction and the intent of the user fully, can carry out intent clarification of its own accord, has a stronger reasoning capability and can complete task decomposition and task orchestration automatically; it can select and call tools dynamically and automatically; it can generate content in at least one modality, text, audio or image; and it has both short-term and long-term memory.

6.4 L4 collaboration level. To be determined. Note: according to the present state of technical development, this document gives the key intelligence capabilities for levels L1 to L3; the division of L4 and higher levels will be made clear and completed in a later revision of this document.

7 Capability elements

7.1 First-level capability elements. 7.1.1 Perception: the capability of the terminal to obtain data from sensors, system services and application services through internal and external sensors, data acquisition modules and recognition modules, and to build the scenario context information the task needs; it comprises perception of user information, of device information and of environmental information. 7.1.2 Cognition: the capability to understand the intent of the user and the related information, to reason about and analyse that information and to plan the task dynamically, so that its execution matches the goal the user expects; it comprises understanding, reasoning and planning. 7.1.3 Execution: the capability to generate a response from the output of the cognitive process and to call internal and external tools and services to achieve the goal of the task; it comprises tool calling, content generation, interconnected collaboration and expressive output. 7.1.4 Memory: the capability, with the user's authorization, to extract, store, retrieve and dynamically update the content of the interaction and the related information; it comprises short-term and long-term memory. 7.1.5 Learning: the capability to improve performance in following instructions, answering knowledge questions and carrying out particular tasks through user feedback, external knowledge input and self-reflection, so as to optimize the output and raise the level of intelligence; it comprises context-adaptive learning and continuous evolutionary learning.

7.2 Second-level capability elements. 7.2.1 Perception. 7.2.1.1 Perception of user information: the capability of the terminal to perceive information of several kinds relating to the user and to collect and recognize the user's biometric features, input content, behaviour and physiological state. Biometric features include voiceprint, face and fingerprint; input content includes text, images, audio and video; behaviour includes gesture, eye movement, remaining still and walking; physiological state includes heart rate and body temperature.