

ICS 35.160
CCS L 62



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

GB/Z 177.3-2026

**Intelligence grading of artificial intelligence terminal - Part 3:
Mobile terminal**

人工智能终端智能化分级 第3部分：移动终端

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	III
Introduction	IV
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 Abbreviated terms	2
5 Key capabilities	2
5.1 Overview	2
5.2 Device-side capabilities	2
5.3 Device-cloud collaborative capabilities	5
6 Determination of the level	10
Annex A (normative) Test methods	11
A.1 General	11
A.2 Test methods for the device-side capabilities	11
A.3 Test methods for the device-cloud collaborative capabilities	17
Annex B (informative) Examples of typical test scenarios	23
B.1 Device-side scenarios	23
B.2 Device-cloud collaborative scenarios	25
Annex C (informative) Examples of tool-calling tasks for a general mobile terminal and their correspondence with difficulty	30
Bibliography	37

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 3 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 provides capability guidance for grading the intelligence of artificial intelligence mobile terminals, referred to below as the mobile terminal, and describes the corresponding test methods.

This document applies to guiding the grading of intelligence of general mobile terminals, and provides a reference for the design, development, application, selection and testing of specialized mobile terminals.

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 of the referenced document, including any amendments, applies.

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

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

GB/Z 177.2-2026, Grading of intelligence for artificial intelligence terminals - Part 2: General requirements

3 Terms and definitions

For the purposes of this document, the terms and definitions given in GB/T 45288.2-2025, GB/Z 177.1 and GB/Z 177.2-2026 and the following apply.

3.1 general mobile terminal - a mobile terminal device on which tools of many kinds can be installed, covering social entertainment, everyday shopping, transport and travel and professional office work, and which can be used in general scenarios. Note: it comprises mainly mobile communication terminals and tablet computers.

3.2 specialized mobile terminal - a mobile terminal device on which only specialized tools can be installed and which is used in a specialized industry field and in a particular scenario.

3.3 enforce Recall-Oriented Understudy for Gisting Evaluation, enROUGE - a summary quality evaluation indicator which, on the basis of the conventional ROUGE indicator, multiplies by a semantic judgement coefficient calculated from semantic similarity, so as to ensure at the same time that the content of the summary is semantically consistent in

direction with the original text and that the ROUGE value is high.

3.4 enforce Bilingual Evaluation Understudy, enBLEU - a translation quality evaluation indicator which, on the basis of the conventional BLEU indicator, multiplies by a semantic judgement coefficient calculated from semantic similarity, so as to ensure at the same time that the translated content is semantically consistent in direction with the original text and that the BLEU value is high.

4 Abbreviated terms

For the purposes of this document the following abbreviated terms apply. BLEU: Bilingual Evaluation Understudy. MB: Mega Byte. ROUGE: Recall-Oriented Understudy for Gisting Evaluation.

5 Key capabilities

5.1 Overview. According to the level of intelligence with which a mobile terminal completes tasks of various kinds within its functional range, and the degree to which the user takes part while the task is carried out, the intelligence level of a mobile terminal is divided into levels 1 (L1) to 3 (L3); the higher the level, the higher the level of intelligence. The capabilities of a higher level cover all the capabilities of the lower levels, so that an L3 mobile terminal covers at the same time all the key capabilities of L1 and of L2. The capability elements of this document follow GB/Z 177.1 and GB/Z 177.2-2026. The intelligence level of a mobile terminal is made up of two parts: a) the device-side capability, meaning the service capability of the mobile terminal to complete part of the perception, cognition, execution, memory and learning locally, relying on its own computing, storage and sensor resources; b) the device-cloud collaborative capability, meaning the capability of the mobile terminal, on the basis of its device-side capability and together with edge-side or cloud-side computing resources, to complete composite tasks that call for strong cognition and complex decision processing, through collaborative scheduling and exchange of information.

5.2 Device-side capabilities. 5.2.1 L1 response level. 5.2.1.1 Perception. 5.2.1.1.1 Perception of user information: the terminal shall be able to perceive the content the user inputs, such as text, images, audio and video. 5.2.1.1.2 Perception of device information: a) it shall be able to perceive the hardware state of the terminal, such as the processor, the memory and the external communication interfaces; b) it shall be able to perceive the software information of the terminal, such as the list of installed software, the permission settings and the running state. 5.2.1.1.3 Perception of environmental information: no requirement. 5.2.1.2 Cognition. 5.2.1.2.1 Understanding: the terminal shall be able to understand simple instructions from the user. 5.2.1.2.2 Reasoning: no requirement.

5.2.3.2 Cognition at L3. 5.2.3.2.1 Understanding: the terminal shall be able to understand

simple instructions from the user. 5.2.3.2.2 Reasoning: no requirement. 5.2.3.2.3 Planning: no requirement. 5.2.3.3 Execution. 5.2.3.3.1 Tool calling: a) the terminal shall be able to call a determinate tool to complete a single-step task; b) it shall be able to call a determinate combination of tools to complete a preset and clearly defined multi-step task. 5.2.3.3.2 Content generation: the terminal shall be able to generate summary or translation content meeting the needs of the user, at an output speed greater than 20 characters per second, with a summary enROUGE-L not less than 0.40 or a translation enBLEU not less than 0.40. 5.2.3.3.3 Interconnected collaboration: no requirement. 5.2.3.3.4 Expressive output: the terminal shall be able to feed back to the user the state and the result of the task through output in the form of images, audio or text. 5.2.3.4 Memory: no requirement. 5.2.3.5 Learning: no requirement.

5.3 Device-cloud collaborative capabilities. 5.3.1 L1 response level. 5.3.1.1 Perception. 5.3.1.1.1 Perception of user information: the terminal shall be able to perceive the content the user inputs, such as text, images, audio and video. 5.3.1.1.2 Perception of device information: the capability covers the hardware state and the software information of the terminal as for the device-side case.

6 Determination of the level

The intelligence level of a mobile terminal is determined by the conformity method of GB/Z 177.2-2026. According to the target intelligence level applied for by the client and the key capabilities this document specifies for that level, not fewer than three typical test scenarios and not less than 80 % of the tool-calling tasks of the difficulty corresponding to that level are selected for scenario-based testing, covering all the key capability requirements. Each scenario verifies the key capabilities associated with it; where some key capabilities cannot be covered, they are tested by the test method for the corresponding capability, the reference test methods being those of Annex A. Examples of typical test scenarios, and the key capabilities each example covers, are given in Annex B; examples of tool-calling tasks for a general mobile terminal and their correspondence with difficulty are given in Annex C.

Where the client considers that the scenarios of this document do not match the positioning of the product, test scenarios may be drawn up independently by reference to Annex B; scenarios drawn up in this way shall be reviewed by the testing body and shall at the same time cover all the capabilities of the corresponding level.