Global Educational Robots Blue Book 2026

Learning in the Age of Foundation Models, Embodied AI, and Humanoid Robots

AI-SLI

June 2026

A Note on Method: How This Report Was Produced with AI Assistance

AI-SLI · Produced with AI Assistance

This blue book is a systematic exploration by AI-SLI into AI-assisted knowledge production, and the 2026 edition of the Global Educational Robots White Paper series (2016, 2019), now upgraded to a blue-book format. Throughout its development, generative artificial intelligence served as a research assistant under the direction and oversight of the research team, undertaking four kinds of labour-intensive work. First, whole-picture collection and data verification: fanning out exhaustive multi-source searches across the market reports, shipment and vendor data, primary policy texts, and scholarly evidence on educational robots, embodied AI, humanoid robots, and agentic tutors; tracing every key figure to its source, corroborating it with no fewer than two independent sources, issuing caliber corrections, and rigorously distinguishing systems that are deployed at scale from those that remain laboratory demonstrations. Second, theoretical anchoring and literature synthesis: extracting and verifying the earlier editions' definitions, taxonomy, and value-chain framework for educational robots from the 2016/2019 editions, and redefining them for the era of foundation models and embodied AI. Third, drafting, figure generation, and citation management: composing the chapters and rendering the architecture and value-chain figures to a unified specification, and maintaining a fully traceable citation apparatus with graded credibility annotations. Fourth, a parallel bilingual edition under a controlled terminology glossary.

We state plainly that the role of AI here was the labour-intensive work of collecting, verifying, drafting, illustrating, and managing citations; the choice of subject, the judgements of value, the scholarly and industry assessments, and the final conclusions were directed and vouched for by the research team. The report holds to conservative accuracy — every datum is required to be real and verifiable, every market figure corroborated by no fewer than two independent sources, and nascent segments are flagged by caliber rather than overstated. We offer it as a forward-looking reference workflow for colleagues in education to scrutinise and improve upon — a sincere experiment in a new paradigm of knowledge production, and in no way a substitute for expert judgement or peer review.

Table of Contents

Part I Overview and Coordinates **1**

Chapter 1 Introduction: A Decade of Educational Robotics and the 2026 Landscape 2

Chapter 2 The Policy Landscape: A Dual-Track Structure Across Twelve Jurisdictions 9

Chapter 3 Definitions and Taxonomy: From a Binary Framework to an Embodied–Agentic Continuum 16

Part II The Capability Leap: Brain, Body, and World Models **24**

Chapter 4 Technical Evolution: From Rule-Based Scripts to Embodied Foundation Models 25

Chapter 5 Agentification: When the Software Tutor Reaches Scale Before the Robot Body 36

Chapter 6 The Full Spectrum of Embodied Bodies: Humanoids, Dexterous Hands, and Quadrupeds 43

Chapter 7 Multimodal Interaction and Learning Analytics: Behavior Sensing, Controlled Research, and the Emotion-Recognition Red Line 55

Part III Deployment and Evidence: Who Actually Reached the Classroom **62**

Chapter 8 The Full Spectrum of Application Scenarios 63

Chapter 9 Cases and Pilots: An Evidence-Graded Map of Real-World Deployment 70

Chapter 10 Learning Outcomes and Evidence: Scenario-by-Scenario Strength Grading 82

Part IV Industry and Market **93**

Chapter 11 Value-Chain Reconstruction: From the Hardware Chain to the VLA Layer and Upstream Bottlenecks 94

Chapter 12 Market Size and TAM: Conservative, Multi-Source Ranges 102

Chapter 13 Competitive Landscape: Representative Vendors, Shipment Rankings, and Capital Divergence 111

Part V Governance and Outlook **120**

Chapter 14 Safety, Ethics, and Governance: Five Regulatory Red Lines and the Hard Constraints on Child–AI Interaction 121

Chapter 15 Trend Assessment, Roadmap, and Policy Agenda 130

Appendices Methods, Glossary, Policy, and Data **137**

Appendix A Research Methodology: Research Objectives, Mixed-Methods Evidence Synthesis, and the Data-Collection Protocol 138

Appendix B Glossary (English–Chinese) 145

Appendix C International Policy Comparison (9 Countries/Regions + 3 International Organizations) 151

Appendix D Chronology of Chinese Policies (2017 → 2026) 158
Appendix E Representative Product Matrix: Model × Capability × Price × Education-Entry Status 164
Appendix F Data Availability Statement and Open-Items List 173
Appendix G References and Data Sources 179

Part I Overview and Coordinates

A single core judgment — the brain has been upgraded, the body has not: the generational leap in capability and the lag in deployment

This part sets out the volume's core judgment and coordinate system: in 2026 educational robotics has crossed a generational threshold in capability, yet deployment lags far behind — “the brain has been upgraded; the body has not.” Before the technical and industrial detail, the policy dual-track and the definitions frame the boundaries of everything that follows.

Chapter 1 Introduction: A Decade of Educational Robotics and the 2026 Landscape

> *When you place "robotics" inside "education," what has happened over the past decade is not incremental refinement but a replacement of the foundation: the "body" of the terminal device has changed only modestly in form, while the "brain" that drives it has been swapped out almost in its entirety. This chapter sets out the core judgment that runs through the whole volume — capability has outrun deployment — and on that basis establishes three narrative spines, five governing disciplines, and a reading map. If you read only one section, read §1.2 on the capability–deployment gap.*

1.1 The Core Judgment: Capability Has Crossed a Generational Threshold; Deployment Has Not Kept Up

If the educational robot of 2026 had to be captured in a single sentence, it would be this: **the most advanced capabilities have crossed a generational threshold, while genuine deployment into routine classrooms lags far behind.** This "capability–deployment gap" is the starting point of every judgment in this blue book, and the root of its divergence from most of the optimistic narratives on the market.

Begin with the "capability" side, for educational robots are not a new concept. In 2016 China's first systematic *Global Educational Robots White Paper* was released, gathering for the first time "robotics education" and "educational service robots" under the single umbrella concept of **Educational Robots** and arguing, with foresight, that Educational Robots would become the "third category of robots" — a development field following industrial robots and service robots [1]. A subsequent revision consolidated the framework into a binary structure, a three-role coordinate system, and a layered value chain, and offered a quantified forecast of the global market [2]. Several independent third parties have reproduced this definition–role–value-chain structure verbatim, and the present research cross-checked it word for word and **found no framework element that could not be traced to a source** [3]. It remains the most systematic theoretical foundation for discussing Educational Robots in the Chinese-language literature, and the point from which this blue book continues and upgrades.

What that generation of educational robots inhabited, however, was in essence a rule-, sensor-, and script-based world. The robot was either a building-block kit that learners assembled themselves and programmed in a visual language, or a fixed-structure service body that ran on a pre-authored action library and scripted dialogue. The "intelligence" of the robot came overwhelmingly from rules an engineer had written in advance, not from the machine's own capacity for language understanding and reasoning.

Within only a few years, the technological substrate turned over at a generational scale. Since 2022, **foundation models** — led by large language models (LLMs) and vision–language models (VLMs) — have given robots the commonsense reasoning, language-based planning, and zero-shot task decomposition that earlier systems lacked, taking on the role of the "brain"; in parallel, end-to-end **vision–language–action (VLA) models** — networks that map visual input directly to motor and joint-control output — have unified perception and action within a single network, acting as the "cerebellum" that coordinates the body [4,5,6]. From NVIDIA's humanoid foundation model GR00T to Figure AI's humanoid VLA system Helix, "a foundation model as the brain plus embodied intelligence as the cerebellum" has become the dominant paradigm in robotics by 2026 [7,8]. The discontinuity is measurable: the list of key technologies for educational robots a decade ago — speech recognition, machine vision, context awareness — and the 2026 technical vocabulary — multimodal large models, embodied intelligence, agent orchestration, on-device inference — barely overlap any longer.

This **paradigm shift** is what renders the older technological narrative, the older market framing, and even the older governance framework inadequate to describe the present. The theory has been updated to match: drawing on generative AI, Huang Ronghuai, Chen Ying, and Ahmed Tlili (2024) reorganized the typical application scenarios of Educational Robots, brought chatbots — ChatGPT among them — explicitly within the scope of the discussion, and proposed four principles of trustworthy AI for education: robustness, legality, compliance, and ethical conformity [9]. Because that update comes from the original author team, the 2026 redefinition has an authoritative and continuous point of connection to it; this blue book is therefore a work of continuation standing on an existing theoretical lineage, not a fresh start.

A leap in capability is not the same as realized deployment. While recording the genuine progress of the paradigm shift, this blue book judges soberly which advances have actually entered the classroom and which remain at the demonstration stage; it embraces the policy dividend of "AI plus education" while confronting the dense wave of child–AI regulatory constraints that landed across 2025 and 2026. The three spines that follow rest on that distinction.

1.2 Three Spines: The Paradigm Shift, the Capability–Deployment Gap, and the Governance Redlines

Three narrative spines run throughout the volume and hold its five parts — landscape, technology, application, industry, and governance and the future — consistent with one another and free of contradiction or overclaim.

Spine one · the paradigm shift. The technological substrate of educational robots has shifted wholesale from "rule, sensor, and script" to "a foundation model as the brain plus embodied-intelligence VLA as the cerebellum." This generational break has clear supporting samples: from the foundational RT-2 to the open-weights π0, and on to the humanoid VLA system Helix and the humanoid foundation model GR00T, robots have for the first time acquired end-to-end language–vision–action capability [4,6,8,7]. Part Two (technology) develops this spine.

Spine two · the capability–deployment gap. This is the most important and most repeatedly stressed of the volume's prudent judgments: in 2026 the most advanced "foundation model plus embodied intelligence" capabilities **remain, for the overwhelming majority, in laboratory demonstrations and vendor release videos, and have not entered routine educational deployment.** In education, what has actually achieved deployment at scale is not the embodied robot body but the on-screen software LLM agentic tutor — Khan Academy's Khanmigo, whose K-12 student users grew from roughly 40,000 to roughly 700,000 within a single school year, while its total users (including teachers) reached roughly 1.4 million by April 2025, covering some 350 U.S. school districts and expanding to India, Brazil, and the Philippines [10,11]. In other words, **the "brain" has landed; the embodied "body" has not.** This blue book requires that every reference to a frontier capability carry a maturity tag (see §1.4), so that a vendor demonstration is never mistaken for an accomplished fact.

Spine three · the governance redlines. Educational robots sit at the intersection of two regulatory forces — one encouraging entry into the classroom, the other constraining the risks of child–AI interaction. China, India, Singapore, the United Kingdom, Japan, Australia, and other jurisdictions are advancing AI and robotics into basic education in curricular form, while across 2025 and 2026 multiple jurisdictions issued hard constraints in quick succession — China's Interim Measures for the Administration of Anthropomorphic AI Interaction Services, the EU AI Act's Article 5(1)(f) prohibition on emotion inference in education institutions, and the U.S. COPPA 2025 amendment bringing children's biometric data under regulation [12,13,14]. This blue book therefore **raises governance from a footnote in the outlook chapters to a first-order framework on a par with definition and taxonomy**, its theoretical point of connection being the four principles of trustworthy AI proposed by Huang and colleagues (2024) [9].

These three spines constrain one another: the paradigm shift explains why "capability" has surged; the capability–deployment gap reminds the reader that "deployment" lags far behind; and the governance redlines define how far deployment may go. **Any claim that stands on its own while detaching from one of the three should be treated as suspect** — this is both the methodological self-discipline of this blue book and a yardstick offered to the reader.

1.3 The Whole Picture and the Structure of the Volume

The system structure of the contemporary educational robot can be summarized as five mutually coupled domains: "brain–body–perception–data–governance." The "brain," composed of foundation models and VLA, supplies reasoning and decision-making; the "body," in social or humanoid form, carries interaction; the multimodal perception layer captures speech, vision, and behavioral signals; the data and platform layer supports training, memory, and operations; and the governance layer imposes external constraints through trustworthy-AI principles, the protection of minors' data, and regulatory redlines. sets out this overview. It also marks the "brain" as deployed at scale (in the form of software agents) and most of the "body" layer as still at the demonstration and pilot stage — the capability–deployment gap is embodied in exactly this mismatch of maturity between the two layers. Fig. 1 is a **structural diagram,**

not a promise of maturity: an embodied form's appearance in it does not mean wide classroom entry, and its specific maturity must be read from the tags on each system in Part Two.

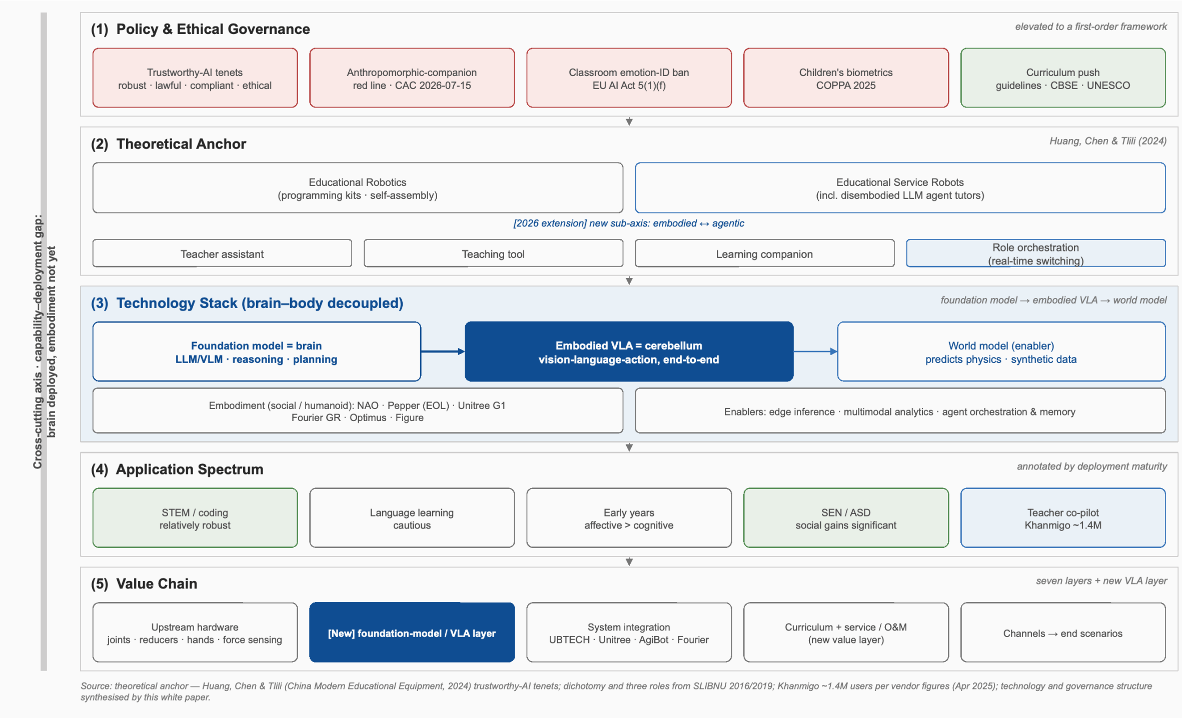


Figure 01.

This blue book comprises fifteen chapters in five parts, plus seven appendices (methodology, glossary, an international-policy comparison, a China policy timeline, a product matrix, data availability, and references). The logical structure of the five parts is shown in .

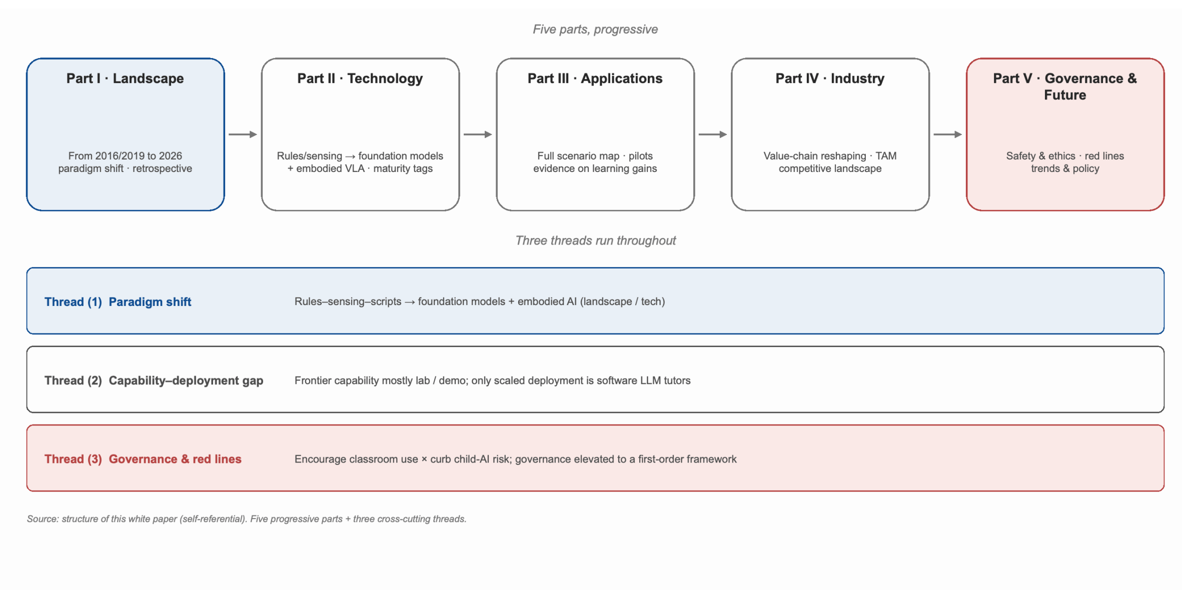


Figure 02.

- **Part One · Landscape (Chapters 1–3):** after this introduction, Chapter 2 gives a survey of the policy landscape across twelve jurisdictions, and Chapter 3 completes the 2026 redefinition of definitions, taxonomy, and educational roles on the foundation of the framework of Huang and colleagues (2024). This is the theoretical and coordinate bedrock of the volume. - **Part Two · Technology (Chapters 4–7):** the four steps of technical evolution, the rise of the agentic tutor, the hardware curve of humanoid and social bodies, and multimodal interaction and learning analytics. Every frontier capability carries a maturity tag. - **Part Three · Application (Chapters 8–10):** the full spectrum of application scenarios; cases and pilots (distinguishing "deployed" from "demonstrated/claimed"); and the evidence base for learning outcomes (with evidence strength flagged by setting, and adverse and null effects recorded on an equal footing). - **Part Four · Industry (Chapters 11–13):** the restructuring of the value chain and the newly added VLA algorithm layer; market size and TAM (presented side by side from multiple sources, with a historical retrospective); and the competitive landscape and the "ice and fire" of capital. - **Part Five · Governance and the Future (Chapters 14–15):** safety, ethics, and governance as a first-order framework (five policy redlines), together with trends, a roadmap, and a research agenda.

Reading guide. Readers seeking a rapid grasp of the whole may read Chapter 1 and each chapter's "summary"; those focused on product deployment should concentrate on Chapters 9 and 13 and Appendix E; those focused on evidence will find Chapter 10 the most information-dense core; and those focused on compliance will find Chapters 2 and 14 and Appendices C and D a complete policy trail. This blue book deduplicates cross-chapter themes explicitly: definitions and redefinition belong to Chapter 3; detailed policy provisions to Chapter 14 and the appendices; market calibers and the historical retrospective to Chapter 12. This introduction only frames the themes and previews them; it does not repeat the arguments.

1.4 Methodology and Five Governing Disciplines

All of the data in this blue book comes from five parallel research streams — market and industry, technical evolution, policy and governance, the evidence base, and the existing theoretical framework and theoretical anchoring — compiled into a single source of fact, with every headline figure cross-verified across multiple sources (Appendix A details the methodology and the credibility grading). Verification follows a graded principle: for every key figure that enters the body text, **the official primary text or a caliber cross-checked by independent third parties is preferred; a single-institution figure must be explicitly flagged; and conflicting calibers are presented side by side rather than resolved in favor of the most aggressive.**

This blue book rests on **mixed-methods evidence synthesis**: a systematic literature review, market-data triangulation, policy-text analysis, and evidence synthesis run in parallel and corroborate one another. A single method would not suffice, because educational robotics is an industry–scholarship hybrid whose market and industry data reside mainly in institutional reports and financial filings rather than peer-reviewed literature. **Data collection** spans four source classes — academic (Google Scholar, ERIC, Web of Science, CNKI), market (institutional reports and industry databases), policy (the official websites of national governments and international organizations), and industry (vendor websites, financial reports, and prospectuses) — over a 2025–2026 collection window, with cross-time figures annotated by collection date, primary distinguished from secondary, and untraceable items removed. **Verification** requires at least two independent sources for each key figure; **credibility** is layered in four tiers — peer-reviewed literature > primary policy texts > institutional reports/financial filings (caliber required) > industry media/secondary aggregation (use with care); and **maturity** runs through the whole volume in four tags — deployed at scale / piloted / lab-demonstrated / vendor-claimed. Appendix A gives the full research-workflow diagram and the complete protocol.

To discipline its writing and citation, this blue book sets itself five governing rules.

1. **The maturity-tag rule.** All technical capabilities are sorted into four levels — "deployed / piloted / demonstrated / vendor-claimed" — and any citation of a frontier capability must carry a tag. "Deployed" means there is public commercial supply or third-party field reporting; "piloted" means small-scale use in a real setting; "demonstrated" means a research paper, a controlled evaluation, or a vendor release video; "vendor-claimed" means a company's own statement that has not been independently verified.

2. **The currency-caliber rule.** Wherever the financials of Chinese companies are concerned, "亿元" and "billion" are aligned strictly to the source-verified value and must not be confused (1 亿元 ≈ US$14 million, certainly not 1 billion). One prominent correction: overseas search engines once misread UBTech's revenue of about RMB 1.305 billion as "13 billion dollars" and Unitree's revenue as "17 billion" — the correct values are RMB 1.305 billion and RMB 1.708 billion respectively (these figures are detailed in Chapter 13).

3. **The evidence-honesty rule.** No effect sizes are fabricated; adverse and null effects are recorded on an equal footing; positive effects observed in a single or short-term session must be discounted for the "novelty effect." For language learning and LLM-driven robots, **no claim of "proven effective" may be made.**

4. **The market-parallelism rule.** Conflicting market calibers are presented side by side with an explanation of the difference, rather than resolved in favor of the single most aggressive value. For example, China's broad-scope market (including training services) and its narrow-scope market (bodies/solutions) differ by nearly threefold; **both must be given at once and never mixed.**

5. **The official-text rule.** A policy's document title, document number, effective date, and key provisions all follow the official primary text; a proposal that has not yet been finally adopted must have its status explicitly flagged.

These five rules are not decorative declarations but hard constraints that run through every figure and every claim in the volume. If the reader finds, anywhere in the book, a frontier capability cited without a maturity tag or a single market figure treated as the sole authoritative caliber, it may be taken as a breach of these rules.

1.5 An Honest Retrospective on Historical Forecasts (Preview)

One methodological improvement of this blue book is an honest retrospective on historical market forecasts. Earlier work, summing four models — consumer terminals, educational institutions, educational kits, and STEAM toys — forecast that the global educational-robot market would reach about **US$11.1 billion** by 2021 and about **US$84.1 billion** by 2023 [12]. Both were **forecasts**, drawn on a broad caliber that folded in large volumes of consumer toys and training services.

The actual 2026 sizing falls far below this. Even on a relatively aggressive institutional caliber, the global educational-robot market reaches only the order of US$5.5–5.8 billion around 2030 [15][16] — which makes the historical forecast of "US$84.1 billion" **plainly too high and too broadly scoped.**

This section is preview only. The multi-source side-by-side ranges for the global and China markets, the item-by-item explanation of caliber conflicts, and the full argument behind the "US$84.1 billion vs. ≤US$5.8 billion" comparison are all presented in Chapter 12. Until then, readers should **strictly distinguish the narrow market for educational-robot bodies from the broad market for robotics education / STEAM consumption**, and treat "US$84.1 billion" as a historical forecast in need of honest correction rather than an accomplished fact to be reused. This retrospective is itself the methodological posture this blue book hopes to convey: **enthusiasm for frontier capabilities, restraint toward numbers and timetables.**

Chapter Summary

This chapter has set out the core judgment that runs through the whole volume — capability has crossed a generational threshold while deployment lags far behind; established three spines (the paradigm shift, the capability–deployment gap, and the governance redlines) and five governing rules (maturity, currency caliber, evidence, parallelism, and official text); given the whole-picture architecture of the five domains "brain–body–perception–data–governance" (Fig. 1) and the five-part reading map (Fig. 2); and previewed the honest retrospective on historical forecasts.

The keynote running through the chapter is prudence, not pessimism: **the "brain" has landed; the embodied "body" has not.** This capability–deployment gap is the premise for understanding every technical, applied, and industrial judgment that follows. Chapter 2 next enters through the policy landscape, surveying the dual-track posture of "encouraging the classroom" and "constraining child–AI" across twelve jurisdictions worldwide, laying the external-constraint groundwork for the redefinition of Chapter 3 and the first-order governance framework of Chapter 14.

Chapter 2 The Policy Landscape: A Dual-Track Structure Across Twelve Jurisdictions

No judgment about educational robots in 2026 can be made apart from the policy field in which they operate. This chapter offers a landscape-level overview of 12 jurisdictions worldwide (9 countries and regions plus 3 international organizations), sketching two parallel tracks — "curricularization in the classroom" and "hard constraints on children–AI interaction" — and the trend by which they converge. One caveat must be stated at the outset: this chapter is a landscape overview only, and all document numbers, effective dates, and clause-level details follow the official originals presented in Chapter 14 and Appendices C and D; this chapter does not unpack the clauses themselves.

2.1 The Dual-Track Structure: Mapping 12 Jurisdictions

The 12 jurisdictions covered by this blue book's policy analysis comprise 9 countries and regions — China, the European Union, the United States, the United Kingdom, Japan, South Korea, Singapore, Australia, and India — together with 3 international organizations: the United Nations Educational, Scientific and Cultural Organization (UNESCO), the Organisation for Economic Co-operation and Development (OECD), and the World Economic Forum (WEF).

The policies of these jurisdictions sort, broadly, into two parallel tracks. **The curricularization track** brings artificial intelligence and robotics into basic education through national curriculum standards, teaching guidelines, compulsory AI-literacy modules, and even local contact-hour mandates. **The hard-constraint track on children–AI interaction** imposes binding limits, through ministerial regulations, statutory provisions, or state laws, on how children interact with AI systems — particularly anthropomorphic companionship, emotion recognition, and biometric collection.

The key judgment is that **the two tracks do not belong to different countries; more often they coexist within a single jurisdiction** — most clearly in China and the European Union. China advances AI in the classroom through its "AI Plus Education" Action Plan, its general-AI-education guidelines, and local contact-hour mandates, while having issued the world's first ministerial regulation aimed at anthropomorphic companionship. The European Union encourages adoption through national guidelines on generative AI in teaching, while its AI Act bans emotion recognition at the level of educational institutions. This "encourage with one hand, constrain with the other" structure is the single most distinctive feature of the global educational-robot policy landscape in 2026. gives a global policy timeline (2018→2027, curricularization versus hard constraints across both tracks).

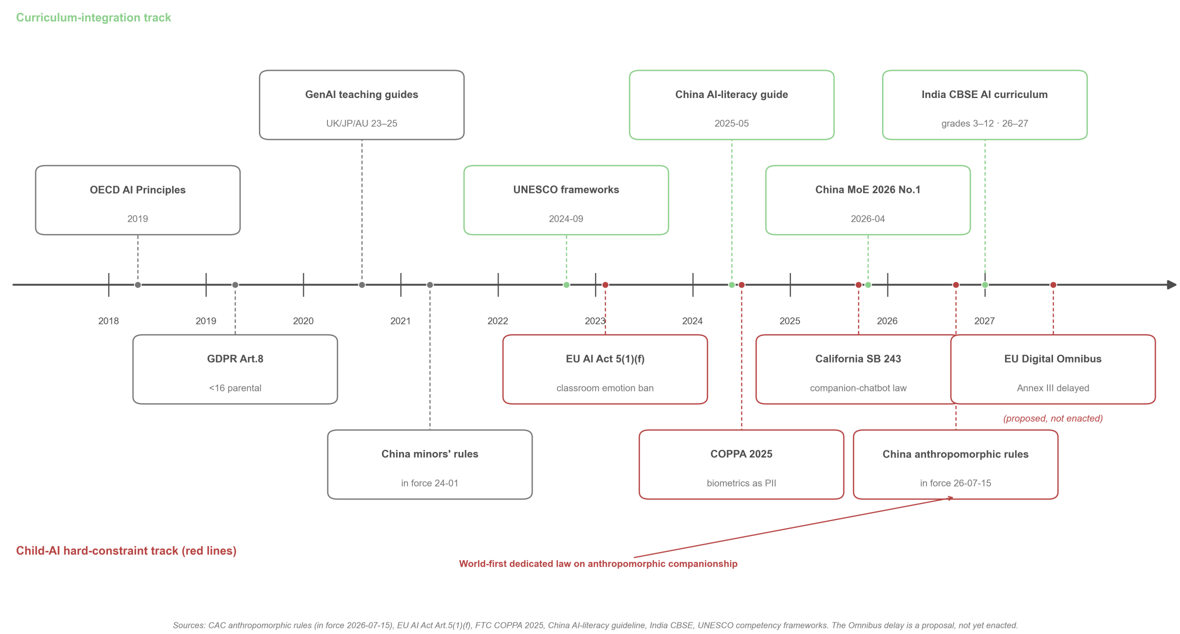


Figure 03.

The three international organizations — UNESCO, the OECD, and the WEF — do not legislate directly, yet through competency frameworks, ethical principles, and agenda-setting they supply a common denominator of values and literacy for national policy, playing a "coordinating-layer" role within the dual-track landscape (see Section 2.5). The survey below treats each jurisdiction along the two tracks before returning to the global judgment of dual-track convergence.

2.2 The Curricularization Camp

The curricularization camp covers China, India, and Singapore, along with the United Kingdom, Japan, Australia, and others. Their common orientation is to elevate AI and robotics from "extracurricular interest" to "curriculum content," and to institutionalize that move through government documents or curriculum standards.

China has pushed curricularization hardest, and its advance has moved from "document-level advocacy" into "contact-hour mandates." At the top level, the State Council's Opinions on Deepening the Implementation of the "AI Plus" Initiative set diffusion targets for intelligent terminals and agent applications [17], while the Master Plan on Building China into a Leading Country in Education (2024–2035) provides the long-range strategic framework [18]. At the implementation level, the "AI Plus Education" Action Plan, jointly issued by five ministries, systematically lays out the path for bringing AI into education and, at the accompanying press briefing, disclosed progress on a national network of AI-education base schools [19][20]. The Guidelines for General AI Education in Primary and Secondary Schools (2025) then set out a tiered, progressive arrangement by stage — experience at the primary level, principles at the junior-secondary level, and systems and innovation at the senior-secondary level [21].

The sharpest instrument is the **contact-hour mandate** at the local level. Beijing's Work Plan for Advancing AI Education in Primary and Secondary Schools (2025–2027) requires that, from the autumn 2025 semester, every primary and secondary school in the city offer a general-AI course of no fewer than 8 hours per academic year, covering all grade levels [22]. Shanghai, from autumn 2024, introduced the local course "Foundations of Artificial Intelligence" in the fourth grade of primary school and the seventh grade of junior secondary school, integrating it into the regular teaching rhythm (media accounts report one period per week and no fewer than 30 hours per academic year; because that hour count is a secondhand report, the original text of the Shanghai Municipal Education Commission should be reviewed before citing a hard figure) [23]. These contact-hour requirements move AI education from "encouraged to offer" to "must be offered in full," and represent the most concrete step the curricularization track has taken in China.

India's Central Board of Secondary Education (CBSE) plans to introduce an AI and computational-thinking curriculum systematically into basic education from the lower grades, with broad coverage and a wide grade span — one of the largest national-level AI-curriculum deployments anywhere [24]. **Singapore** has folded AI literacy into its education-technology plan oriented toward 2030 and set a compulsory AI-literacy module, emphasizing digital and AI competence for all [25].

The United Kingdom, Japan, and Australia rely chiefly on teacher-led guidelines for generative AI in teaching: the UK Department for Education has issued guidance on generative AI for schools [26]; Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT) has updated its guidelines on the use of generative AI in schools [27]; and Australia has issued a framework for generative AI in education built around the core value of "human and social wellbeing" [28]. What this set of policies shares is a cautious rather than aggressive stance — emphasizing the introduction of AI tools under teacher supervision and toward clear educational goals, rather than leaving students to use them on their own.

The policy dividend of the curricularization camp is clear: products and curricula aligned with general-education guidelines, the CBSE curriculum, and UNESCO competency frameworks find it easier to win direct institutional purchases and public procurement. That dividend always carries the precondition of "teacher-led, fully validated" adoption — a precondition thrown into relief, in reverse, by the cautionary case of South Korea's AI digital textbooks discussed in Chapter 14.

2.3 Industrial Policy: Local Embodied-/Humanoid-Robot Support and Its Link to Education

A second force has shaped the supply side of educational robots between 2025 and 2026: the dense stream of **local-government support policies for embodied intelligence and humanoid robots**. These do not target classrooms directly, but by cultivating robot makers, lowering hardware costs, and supporting research-and-teaching platforms they indirectly determine "what kind of robot can enter education at what price."

At the national level, the Ministry of Industry and Information Technology's 2023 guiding opinions on the innovative development of humanoid robots established the direction of cultivating humanoids as a strategic emerging industry [29]. At the local level, Shenzhen, Shanghai, and Beijing have each issued three-year embodied-intelligence action plans: Shenzhen aims for more than 1,200 embodied-related enterprises and an associated industry scale exceeding RMB 100 billion by 2027 [30]; Shanghai aims for a core embodied-industry scale exceeding RMB 50 billion by 2027, with breakthroughs in a set of core algorithms [31]; and Beijing aims to cultivate no fewer than 50 core enterprises by 2027 and to land embodied robots at the ten-thousand-unit scale, explicitly listing research and education as one of the scenarios for large-scale application [32]. Alongside this industrial momentum, humanoid-robot firms in Shenzhen and elsewhere are accelerating toward the capital markets [33].

The transmission of this industrial policy to education is indirect but real: local support has accelerated the domestic production and cost reduction of robot bodies, bringing quadrupeds and research-and-teaching humanoids into university and vocational laboratories at more affordable prices. A careful distinction is warranted, though — the "large-scale application" of these industrial goals refers chiefly to industrial, commercial, and research settings, and **does not amount to "humanoids entering primary and secondary classrooms as teachers."** The logic linking industrial policy to curriculum policy is developed in the "policy-driven" section of Chapter 11 on the value chain.

2.4 The Children–AI Hard-Constraint Camp

Running parallel to curricularization is the cluster of hard constraints on children–AI interaction that landed densely between 2025 and 2026. Their focus is concentrated in three directions: anthropomorphic companionship, classroom emotion recognition, and children's biometrics. The representative jurisdictions are China, the European Union, and the United States.

China is likewise at the front of the hard-constraint track. The Interim Measures for the Administration of Anthropomorphic AI Interaction Services [unofficial translation; no official English version as of the data cut-off], promulgated on April 10, 2026, and **taking effect on July 15, 2026**, is the most consequential of these: it bans the provision to minors of "virtual kin" and "virtual companion" intimate-relationship services, and prohibits the inducement of emotional dependence and emotional manipulation [12]; the accompanying Q&A further clarifies its scope and regulatory intent [34]. This is the world's first ministerial regulation aimed directly at "anthropomorphic companionship," and it bears especially directly on how companion-type educational robots position their products. In parallel, the Guidelines for the Use of Generative AI in Primary and Secondary Schools (2025) set limits at the instructional level — barring primary-school students from using open-ended content-generation features on their own, and requiring schools to establish a whitelist of AI tools [35].

The European Union's hard constraints are carried by the AI Act (Regulation (EU) 2024/1689) [36]. Its Article 5(1)(f) **prohibits the use of AI to infer the emotions of natural persons in educational institutions** (that is, emotion recognition), applicable from February 2, 2025 [13]; it directly restricts classroom robots and camera systems with "attention- or emotion-monitoring" features, and serves as a global design-prohibition warning for multimodal learning-analytics products (see Chapters 7 and 14).

The United States distributes its hard constraints across the federal and state levels, and state legislation has already moved from "proposal" to "in force." Federally, the 2025 amendment to COPPA brought biometric information (such as voiceprints and facial features) within the scope of children's personal information, taking effect June 23, 2025, with a compliance deadline of April 22, 2026, and constraining educational robots that capture children's voices and faces [14]. At the state level, California's SB 243, the Companion Chatbots Act, was signed in October 2025 and **took effect January 1, 2026**, requiring disclosure of AI identity and the setting of self-harm/suicide crisis-response protocols [37]; New York State's AI-companion law likewise took effect November 5, 2025, requiring crisis intervention and periodic "not a real person" disclosure [38].

"Upstream" of these hard constraints lies a shared data-privacy foundation common to all jurisdictions: the EU's GDPR sets a parental-consent threshold for processing minors' data [39]; the United States constrains education data through FERPA and student-privacy protections [40]; China's Regulations on the Protection of Minors in Cyberspace systematically govern minors' online environment and data [41], while the Interim Measures for the Administration of Generative Artificial Intelligence Services require measures to prevent addiction among minors [42]. This privacy foundation, layered atop the three hard constraints above, constitutes the multiple compliance requirements that any educational robot capturing children's voice, face, and behavioral data must satisfy simultaneously (see the data-privacy-foundation section of Chapter 14).

A calibration note on the EU regulatory timeline. In November 2025, the European Commission proposed a "Digital Omnibus," intended to defer the application of certain standalone high-risk obligations under the AI Act to December 2, 2027. Although that proposal reached a provisional political agreement in May 2026 and was endorsed by a European Parliament vote in June 2026, **as of June 2026 it has not been formally adopted and has not been published in the Official Journal; the legal baseline remains Regulation (EU) 2024/1689** [4336]. One point requires special emphasis: even if the proposal is ultimately adopted, **the Article 5 prohibitions (including the ban on emotion recognition in educational institutions) are unaffected by the deferral and have applied on schedule since February 2, 2025**. Wherever this blue book touches on the proposal, it consistently marks it as "a proposal under provisional agreement, not yet adopted," and refers the clause-level details to Chapter 14 and Appendix C.

2.5 Dual-Track Convergence and the Coordinating Role of International Organizations

Viewing the curricularization track and the hard-constraint track together yields a clear global judgment: **global educational-robot policy is converging from "one-directional encouragement into the classroom" toward a dual-track structure of "encouragement plus hard constraints in parallel."**

Three signatures mark this convergence. **Encouragement and constraint often issue from the same jurisdiction** — both China and the European Union possess strong curricularization policies and binding child-protection regulations at once, showing that the two are not opposed routes but two faces of a single governance logic. **The focus of constraint is highly consistent** — whether in China, the European Union, or the United States, regulation falls on the same three foci: anthropomorphic companionship, classroom emotion recognition, and children's biometrics. And **international organizations supply a common denominator of values and competencies** — UNESCO's AI Competency Frameworks for Students and for Teachers [44] and the OECD's AI Principles [45] give national policies a shared ethical and literacy foundation, making the dual-track convergence amenable to international coordination.

For product makers and educational institutions, this convergence means that compliance can no longer be an after-the-fact remedy but must be moved forward into a seed constraint on product design: companionship features must avoid a "virtual kin / virtual companion" positioning and strengthen "not a real person" disclosure; classroom applications must steer clear of emotion recognition; children's data must be handled so as to satisfy multiple jurisdictions at once; and curriculum content must align with national general-education and literacy frameworks to capture the policy dividend. The full development of these compliance implications appears in Chapter 14.

The boundary of this chapter bears restating. This chapter is a **landscape overview**, intended to build the reader's overall grasp of the "dual-track structure" rather than to recite policy clauses one by one. The document numbers, effective dates, scope of application, and key provisions of each policy are developed, as a first-order framework, in Chapter 14 (safety, ethics, and governance), with Appendix C (a comparative table of international policy) and Appendix D (a timeline of Chinese policy, 2018→2027) supplying a jurisdiction-by-jurisdiction, point-in-time comparison. Readers can follow the thread "Chapter 2 landscape → Chapter 14 clauses → Chapter 11 industrial drivers" to grasp in full how policy shapes the development of educational robots.

Chapter Summary

This chapter has given a landscape overview of educational-robot policy across 12 jurisdictions (9 countries and regions plus 3 international organizations), identifying two parallel tracks — "curricularization in the classroom" and "hard constraints on children–AI interaction" — and offering the global judgment of "dual-track convergence": encouragement and constraint often issue from the same jurisdiction, the focus of constraint is highly concentrated on anthropomorphic companionship, classroom emotion recognition, and children's biometrics, and international organizations supply a common denominator of values and competencies. The chapter has also folded in the indirect transmission of local embodied-/humanoid-robot industrial support to the education supply side, along with the latest 2026 developments — China's local contact-hour mandates and the "in force" status of US state-level companion-robot legislation. A global policy timeline appears in Fig. 3. The chapter holds strictly to its boundary of "landscape overview only, no clause-level unpacking"; all details follow the official originals in Chapter 14 and Appendices C and D. Wherever the EU Digital Omnibus is concerned, it is consistently marked "a proposal under provisional agreement, not yet adopted; the baseline remains Reg. (EU) 2024/1689, and the Article 5 ban on emotion recognition is unaffected."

The policy landscape sets the external boundary of "how far educational robots can go," but to understand what they *are* we must return to the definition itself. The next chapter, building on the dual-category framework and three-role coordinates established by the SLIBNU team and the authoritative update by Huang Ronghuai, Chen Ying, and Tlili (2024), completes a 2026 re-definition of educational robots' definition, taxonomy, and educational roles — the theoretical bedrock of the whole report, and the close of this opening part on the landscape.

Chapter 3 Definitions and Taxonomy: From a Binary Framework to an Embodied–Agentic Continuum

Definitions are this blue book's coordinate bedrock. Building on the dual-category framework and three-role coordinates established by the Global Educational Robots White Paper series, and on the authoritative update by Huang Ronghuai, Chen Ying, and Ahmed Tlili (2024), this chapter completes a 2026 re-definition of educational robots' definition, taxonomy, and educational roles. It preserves the theoretical core that, on verification, holds firm; it adds two new dimensions that capture the paradigm shift; and it elevates governance to a constitutive element of the framework — every new proposal anchored to a citable basis. The three-dimensional coordinate system of the re-definition (dual category × embodied/agentic × role orchestration) appears in .

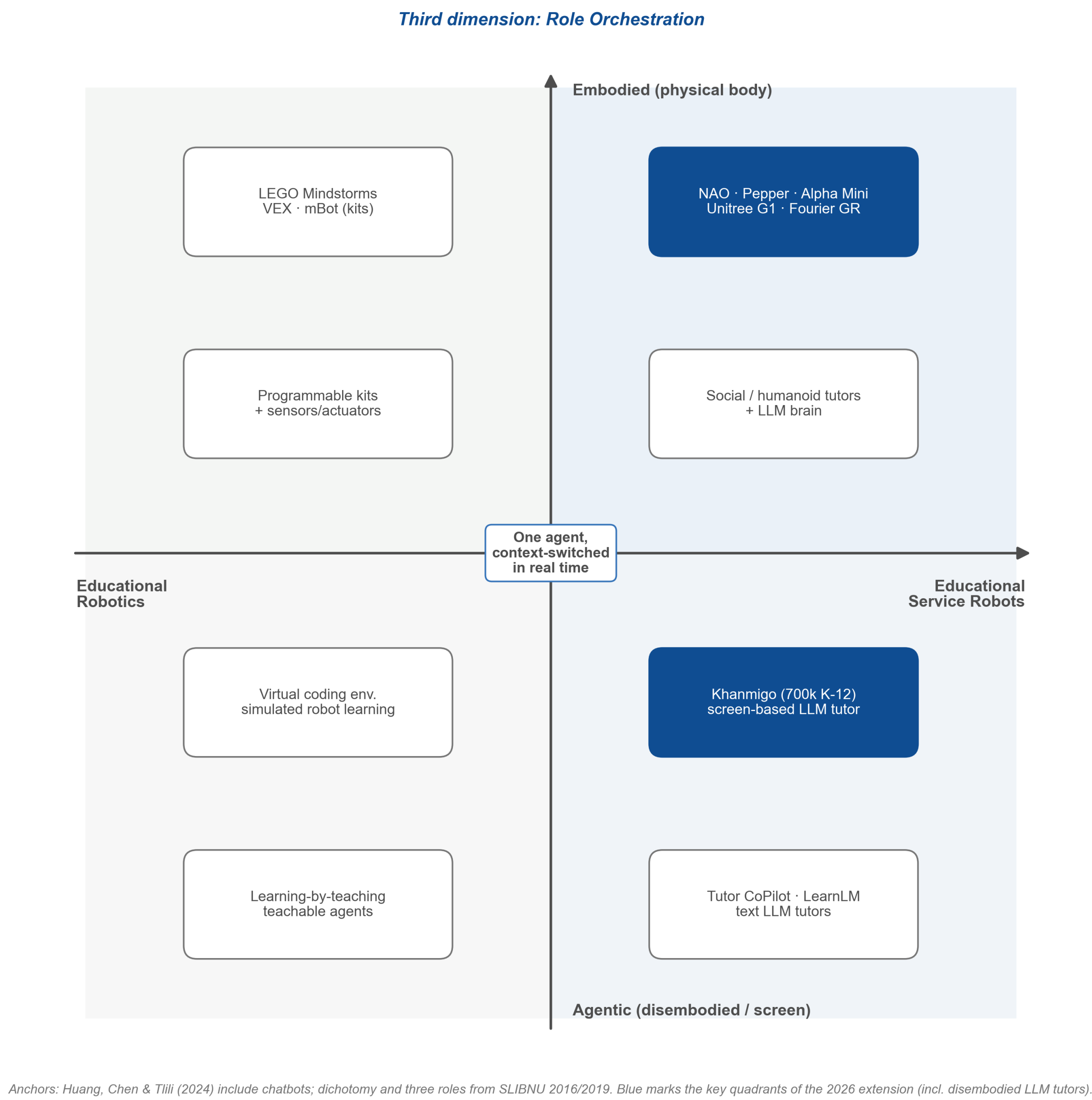


Figure 04.

3.1 The Stable Core: The General Definition, the Binary Framework, and the Three-Role Coordinates

The general definition of educational robots has a valuable inclusiveness and openness: "all robotics for education that assists teaching or learning activities — 'Educational Robotics' — together with service robots that possess educational-service intelligence — 'Educational Service Robots' — are collectively termed Educational Robots." [12] This definition does not require an educational robot to have a humanoid or physical form; its core criteria are "assists teaching or learning" and "possesses educational-service intelligence." For exactly that reason, a decade later it can still consistently accommodate the full spectrum from building-block kits to purely software LLM dialogue tutors — and this openness is the key basis on which the re-definition in Section 3.4 rests.

Beneath the general definition lies the core binary of educational robots [12]:

- **Educational Robotics**: characterized by kit-based learning in which learners assemble and program robots themselves, with typical products such as LEGO Mindstorms and mBot. Its educational value lies in "learning by doing" — learners cultivate computational thinking and engineering literacy by building and programming robots. - **Educational Service Robots**: service robots that possess teaching-and-learning intelligence, with fixed structures that generally do not support learner disassembly and reassembly, applied in scenarios such as STEAM education, language learning, and special education.

This binary is clear and complementary: programmable kits (objects the learner modifies) and intelligent teaching service robots (objects the learner interacts with) represent two fundamentally different educational relationships. This blue book preserves the binary as a base on which to build, and independent third-party reproductions of the framework corroborate its constituent elements [346].

On the instructional dimension, educational robots were defined into three principal roles: **teaching assistant, teaching tool, and learning companion** [2]. These three form the coordinate core for understanding what educational robots "do." The earlier exposition already anticipated that an educational robot might play the multiple role of "both teacher and friend" [1] — a forward-looking observation that lays, within the original framework, the groundwork for this chapter's Section 3.5 upgrade of fixed roles into "orchestratable roles."

Examined against the technological and market environment of 2026, this definitional structure yields one overall judgment: **the core of definition–binary–three roles still holds, and multi-source cross-verification found no constituent element of the framework that could not be traced to a source** — it is a stable base on which to build, not an old framework to be torn down and rebuilt. The base does, however, face three tensions arising from the paradigm shift. The binary accommodates non-physical forms but does not explicitly characterize "degree of embodiment," a dimension that becomes pivotal in the humanoid/VLA era. The three roles anticipated multiple roles yet are organized as "fixed roles" rather than "orchestratable roles." And the key-technology list — speech recognition, machine vision, context awareness — has been superseded by multimodal large models, VLA, and agent orchestration. These three tensions are the entire motivation for the re-definitions in Sections 3.4–3.6: not innovation for its own sake, but the bridging of specific gaps between a stable base and a new paradigm.

3.2 The Primary Anchor: The Authoritative Update to the Taxonomy (2024)

The re-definitions rest on one anchor above all others: the update the original author team completed in 2024. Drawing on generative AI, Huang Ronghuai, Chen Ying, and Ahmed Tlili (2024) extended the educational-robot framework on their own initiative [9], chiefly in three respects:

1. **Mapping 7 typical application scenarios** — spanning STEM educational robots, robotics-oriented programming, social-robot-supported language learning, and more — converging scattered applications into a discussable spectrum of scenarios. 2. **Bringing chatbots into scope** — using ChatGPT as an

example to discuss the value of chatbots in school education, that is, explicitly bringing non-physical LLM agents within the scope of the educational-robot discussion. 3. **Proposing four principles of trustworthy AI** — robustness, legality, compliance, and ethical conformity — elevating governance and ethics from a single principled statement ("designed in accordance with social and ethical norms") into a systematic set of principles.

This update serves as the primary anchor for the chapter's re-definition, rather than an ordinary reference, on two counts.

It provides **public grounds, from the original author team, for bringing LLM agents within the scope of educational robots**. The 2026 edition's judgment — that "Educational Service Robots must encompass non-physical LLM agent tutors" — is thus not an imposition by outside researchers but a continuation of the original team's own direction. This blue book treats a screen-based software agent such as Khanmigo as the core sample of "the only educational application deployed at scale" [10], and its theoretical legitimacy stems from exactly this.

And by elevating governance from a principled statement into a systematic set of principles, it provides **a continuation point, from the original author team**, for this blue book's treatment of governance as a first-order framework (Chapter 14). The four principles of trustworthy AI are therefore the shared anchor of both the definitional re-definition in Chapter 3 and the governance framework in Chapter 14 — the two chapters divide labor without overlapping: this chapter argues why governance belongs in the definition, and Chapter 14 sets out the five specific risk domains of the governance framework.

The three re-definitions that follow are continuations atop this anchor, not departures from scratch. This is the key to understanding the entire chapter.

3.3 The "Third Category of Robots" Thesis and the 2026 Standing of the Historical Product Framework

Beyond the definitional structure, the earlier exposition advanced a forward-looking thesis: that educational robots would become the **"third category of robots"** as a field of development, after **industrial robots** and **service robots** [1] — a thesis corroborated by a verbatim third-party reproduction [3]. Viewed from 2026, the judgment has considerable insight: education has become one of the explicitly discussed downstream application directions for the embodied-intelligence and humanoid-robot industry (though, as this blue book repeatedly stresses, education is its downstream, not its main arena).

On the product framework, the earlier edition derived multiple product forms from a combination of "target user × application setting," and offered a multidimensional product-evaluation framework [12]. The main challenge this framework faces in 2026 is that it struggles to accommodate non-physical forms such as "AI education agent subscriptions" — precisely the problem that the key-technology-list replacement in Section 3.6 addresses.

3.4 Re-definition One: Adding an "Embodied ↔ Agentic" Sub-Dimension Beneath the Binary

The first re-definition adds, beneath the "Educational Robotics ↔ Educational Service Robots" binary, an orthogonal sub-dimension: **"embodied ↔ agentic/disembodied."**

Its **motivation**: while the binary accommodates non-physical forms, it does not explicitly characterize "degree of embodiment." And by 2026 the forms of educational robots have spread out along a continuum — at one end, highly embodied physical bodies such as NAO and humanoid platforms; at the other, purely software LLM dialogue tutors such as Khanmigo; and in between, hybrid forms of "a low-cost social body plus a cloud or on-premises LLM brain" [10]. The binary alone cannot distinguish the essential difference, in degree of embodiment, between "a humanoid that plugs in an LLM brain" and "a purely software agent tutor."

This sub-dimension has two **bases**. First, Huang et al. (2024) have explicitly brought chatbots within the scope of educational robots [9], so the "agentic" end now has the support of the original author team. Second, the general definition itself does not require an educational robot to have a humanoid or physical form [1], so adding a "degree of embodiment" sub-dimension makes the original definition explicit rather than violating it. The maturation of embodied intelligence as an independent technical domain provides technical-side support for this sub-dimension [47].

With this sub-dimension added, the taxonomy of educational robots expands from a "binary" into a two-dimensional grid of "binary × embodied/agentic continuum," letting the definition consistently cover the full spectrum from humanoid bodies to purely software tutors, and providing a shared coordinate for Chapter 5's analysis of "brain ↔ body separation" in agentification and Chapter 6's hardware analysis of humanoid bodies.

3.5 Re-definition Two: Roles Upgraded from Fixed to Orchestratable (Role Orchestration)

The second re-definition restates the "three roles / multiple roles" as **role orchestration**: a single agent can switch in real time among tutor, peer, assessor, and other roles according to the instructional situation, rather than being fixed in a single role.

Its **motivation**: in the script era, roles were "three fixed points," whereas in the large-model era an agent can dynamically switch functional positioning within a single session. In Khanmigo's practice, one software agent can both guide students through Socratic questioning (the tutor role) and generate lesson plans and rubrics for teachers (the assistant role) [10] — exactly how role orchestration manifests in reality.

This upgrade also has **groundwork** from the original framework: the earlier exposition already anticipated that educational robots might take on multiple roles "as both teacher and friend" [1]. This study merely makes that forward-looking observation explicit — upgrading it from "the robot can be one of several roles" to "the robot can orchestrate multiple roles in real time." Beyond the fixed three roles, the orchestratable framework naturally accommodates several new roles — Socratic-questioning tutor, formative assessor, learning-data-analytics advisor, and collaborative-learning facilitator — all functional facets the same agent is orchestrated into across different situations.

Role orchestration upgrades the three-role coordinates from static to dynamic, and is the definitional-side basis for Chapter 5's analysis of "agent orchestration and memory." The division of labor between this chapter and Chapter 5 is: this chapter gives the definitional framework of "orchestratable roles," and Chapter 5 gives its technical implementation (tool calling, memory, sub-agent validation).

3.6 Re-definition Three: Replacing the Technology List and Elevating Governance to a Constitutive Element

3.6.1 Replacing the Key-Technology List

The original key-technology list — represented by speech recognition, machine vision, and context awareness — is no longer sufficient to describe the technical foundation of 2026. This study proposes updating it to: **multimodal large models / embodied intelligence (VLA) / agent orchestration and memory / on-device inference compute / trustworthiness and alignment.**

This replacement has two **bases**. First, Huang et al. (2024) have already brought chatbots and generative AI into the discussion, providing the original author team's basis for adding "multimodal large models" to the list [9]. Second, the maturation of embodied intelligence (VLA) as an independent technical domain provides the technical-side basis for adding "embodied intelligence" [47]. The item-by-item technical evolution of the new list (the four steps from rules and sensing to embodied foundation models) is developed fully in Chapter 4; the division of labor between this chapter and Chapter 4 is: this chapter gives "why the list must be replaced," and Chapter 4 gives the technical detail.

3.6.2 Elevating Governance to a Constitutive Element

In the earlier framework, governance was condensed into a principled sentence, whereas Huang et al. (2024) have already elevated it into the systematic four principles of trustworthy AI [9]. Building on this, the study further proposes listing **protection of minors' data, model hallucination and academic misguidance, emotional dependence, compute equity, and the boundary of human–machine responsibility** as **first-order elements** of the definitional framework, rather than as footnotes to a forward-looking chapter.

This elevation directly continues the four principles of trustworthy AI [9] and resonates with the dense landing of global hard constraints described in Chapter 2. This chapter (the definitional side) argues "why governance should be a first-order element of the definitional framework," and Chapter 14 (the governance side) sets out "the specific clauses and red lines of the five risk domains"; the two divide labor without overlapping. This is the definitional bedrock for the whole report's elevation of governance "from a footnote to a first-order framework."

3.6.3 Industrial and Market-Scope Changes Derived from the Re-definition

Replacing the key-technology list carries one further industrial-side corollary: at the value-chain level, "the foundation-model and embodied-AI algorithm layer (the VLA layer)" becomes an entirely new structural layer. This judgment of an added layer shares a source with the "embodied/agentic sub-dimension" of Section 3.4 (both stem from the maturation of the embodied-intelligence technical domain and the inclusion of chatbots); its full argument belongs to Chapter 11, and this chapter only notes its shared source with the re-definition. Likewise, the market scope must be reworked because of the emergence of non-physical forms (AI education agent subscriptions / MaaS) — the historical four-model summation can no longer accommodate the new forms, and the corresponding parallel scopes and historical retrospect belong to Chapter 12.

3.7 The Overall Coordinate System of the Re-definition

Synthesizing the three re-definitions of Sections 3.4–3.6, the 2026 re-definition can be summarized as a three-dimensional coordinate system (see Fig. 4):

- **Dimension 1 · the binary (inherited from the original framework)**: Educational Robotics ↔ Educational Service Robots; - **Dimension 2 · embodied/agentic (re-definition one)**: embodied body ↔ agentic pure software, forming a continuum; - **Dimension 3 · role orchestration (re-definition two)**: the real-time-orchestration space of tutor / peer / assessor and other roles.

The core of this coordinate system remains the stable binary and three roles — it does not discard the existing framework but adds atop it two dimensions that capture the paradigm shift, and elevates governance to a constitutive element running through all three dimensions. Any educational robot within the coordinate system can be located by three questions: "which of the binary categories does it belong to × how embodied is it × which roles can it orchestrate." Khanmigo is "Educational Service Robot × agentic × orchestratable tutor/assistant"; a NAO-plus-LLM special-education solution is "Educational Service Robot × embodied × orchestratable peer/assessor"; and LEGO Mindstorms is "Educational Robotics × embodied (object the learner modifies) × weak role orchestration." This coordinate system will serve as the shared locating tool for the subsequent chapters on technology, applications, and industry; its term definitions appear in Appendix B.

Placing this re-definition in the broader landscape of smart education reveals its continuity: educational robots are not an isolated technical category but a node — "the embodied agent" — within the smart-education ecosystem, co-evolving with smart learning environments, education large models, and learning analytics [48]. The re-definition coordinate system of this blue book is designed precisely to locate, consistently and within that larger landscape, the full spectrum of educational robots "from purely software tutors to humanoid bodies"; it is backward-compatible with the stable binary base and forward-connected to the overall framework of smart education, so that the theoretical coordinates of educational robots in 2026 are continuous rather than fractured.

Finally, the chapter's methodological discipline bears restating: the three re-definitions above are all proposals by this study extended atop authoritative anchors, each anchored to a citable basis (Huang et al.'s 2024 inclusion of chatbots and the four principles of trustworthy AI, the original definition's not requiring a physical form, the earlier exposition's "both teacher and friend" groundwork, and the maturation of the embodied-intelligence technical domain). When citing this chapter, readers should always be able to distinguish what is the verified, stable core framework from what is this study's extended proposals atop it — this traceability is the concrete embodiment, in the definitions chapter, of this blue book's iron rule of "inventing no untraceable framework."

Chapter Summary

Building on the *Global Educational Robots White Paper* series' binary framework and three-role coordinates, and on the authoritative update by Huang Ronghuai, Chen Ying, and Tlili (2024), this chapter has completed the 2026 re-definition of educational robots:

- **The stable core**: the general definition, the Educational Robotics ↔ Educational Service Robots binary, the teaching-assistant / teaching-tool / learning-companion three roles, and the "third category of robots" thesis — all still hold on multi-source cross-verification. - **The primary anchor**: Huang et al.'s (2024) inclusion of chatbots within scope and their four principles of trustworthy AI are the continuation point of the re-definition, not a departure from scratch. - **Three re-definitions (each anchored to a basis)**: adding an "embodied ↔ agentic" sub-dimension beneath the binary; upgrading roles from fixed to orchestratable; and replacing the key-technology list while elevating governance to a constitutive element.

The three-dimensional coordinate system of the re-definition (binary × embodied/agentic × role orchestration) appears in Fig. 4, and its core remains throughout the stable binary and three roles.

With this coordinate bedrock established, the opening part of this blue book (the landscape) is complete. The second part will develop the technical analysis along the "embodied/agentic continuum": Chapter 4 traces the four technical steps from rules and sensing to embodied foundation models; Chapter 5 focuses on the agentic tutor at the "brain ↔ body separation"; Chapter 6 examines the hardware curves of humanoid and social bodies; and Chapter 7 considers multimodal interaction and learning analytics — every frontier capability carrying the maturity label this blue book requires.

Part II The Capability Leap: Brain, Body, and World Models

Foundation models (the brain) · embodied intelligence / VLA (the cerebellum) · world models (the enabling technology)

Along a “brain–cerebellum–enabler” spine, this part decomposes the 2026 frontier capability stack: foundation models as the brain, embodied-intelligence VLA as the cerebellum, and world models as the enabling technology — and explains why the software “tutor brain” reached scale before the robot body. Every capability carries a maturity tag, so a vendor demo is never mistaken for an accomplished fact.

Chapter 4 Technical Evolution: From Rule-Based Scripts to Embodied Foundation Models

> ***Maturity-label convention*** *(used throughout this part for every technical entry):* ***deployed*** *denotes public commercial supply or independent on-site reporting;* ***piloted*** *denotes small-scale use in real settings;* ***demonstrated*** *denotes a research paper, a controlled evaluation, or a vendor release video;* ***vendor-claimed*** *denotes company figures not independently verified. Every reference to a frontier capability carries one of these tags.*

4.1 Introduction: A Generational Break That Genuinely Exists

Less than a decade ago, the "intelligence" of an educational robot rested chiefly on three families of technology: rule-based deterministic control, sensor-based perception of the environment, and pre-written dialogue and motion scripts. A NAO that "recognized" a student's emotion or "answered" a question in class was, at bottom, advancing along engineer-defined branches inside a finite-state machine. It could do a great many things, but every one of them was hard-coded; the moment it left the script, it was helpless.

The technical picture of 2026 has changed qualitatively. The most advanced systems today rest on a new paradigm of "the foundation model as the brain, embodied intelligence (vision–language–action, VLA) as the cerebellum": large language models and vision–language models supply commonsense reasoning, language planning, and zero-shot task decomposition, while a VLA network maps visual input end-to-end into joint-control output; and above both, world models emerge that can predict physical consequences and generate training data. This leap is **a generational break that genuinely exists** — not the extension of a single technology curve but a change of paradigm (against the contrast between the script-driven era's representative works [1] and [2] and today's [4], [5], and [6], the break is clearly legible).

A second judgment of equal weight rides alongside the first: **a generational leap in capability is not a generational leap in deployment**. The chapter traces the technical thread along "four steps," tags each representative system with a maturity label, and aggregates the distribution in §4.7 — where the great majority of the frequently cited frontier systems prove to remain, to date, at laboratory demonstration and vendor release video rather than routine educational deployment. This **capability–deployment gap** — frontier capabilities lodged in lab demos and vendor videos rather than ordinary classrooms — is the core argument running through every technical chapter.

4.2 The Four-Step Thread: Rules/SLAM → Deep-Learning Perception → Foundation Model as Brain → VLA/World Model as Cerebellum

Compressing the technical evolution of the past decade into a single spine, it can be summarized as four progressive steps (for the survey basis, see [49] and [50]):

Step one · rules, sensing, and SLAM (about 2019 and earlier). This stage is represented by deterministic control, hand-engineered features, and simultaneous localization and mapping (SLAM). Systems were highly capable in structured environments but brittle and hard to generalize: once the environment departed from design assumptions, performance degraded sharply. In education, LEGO Mindstorms, VEX, and early NAO all belong here — their "behavior" was rule-based, predictable, and closed.

Step two · deep-learning perception (about the 2012–2020 period). Convolutional neural networks and Transformers were applied to visual and speech perception, and robots "saw more accurately and heard more clearly." But at this stage **perception and decision-making remained separate**, and tasks were highly specialized: a model that recognized gestures could not transfer to understanding object grasping, and every capability required its own data and training pipeline.

Step three · the foundation model as "brain" (from 2022). Large language models and vision–language models brought commonsense reasoning, natural-language planning, and zero-shot task decomposition. For the first time a robot could break an open-ended instruction down into executable sub-steps (as shown by work such as SayCan and Code as Policies). The first thing truly deployed at scale on this step was not an embodied robot but **a software agent** — the core fact developed in §4.4 and Chapter 5.

Step four · embodied intelligence and VLA as "cerebellum" (from 2022). Vision, language, and action were unified into an end-to-end network that generates action directly from pixels and instructions, bridging the gulf between perception and control; and on this basis, **world models** further evolved — generative systems that can predict the physical consequences of actions and produce synthetic training data. Step three supplies the cognition of "what to do," and step four the motor intelligence of "how to do it"; the combination of the two defines the frontier of 2026.

The four steps **layer rather than replace** one another in time. What runs at scale in classrooms is still, to this day, chiefly the programmable kits of step one; the capabilities of steps three and four, though repeatedly demonstrated, have not yet settled into routine educational products.

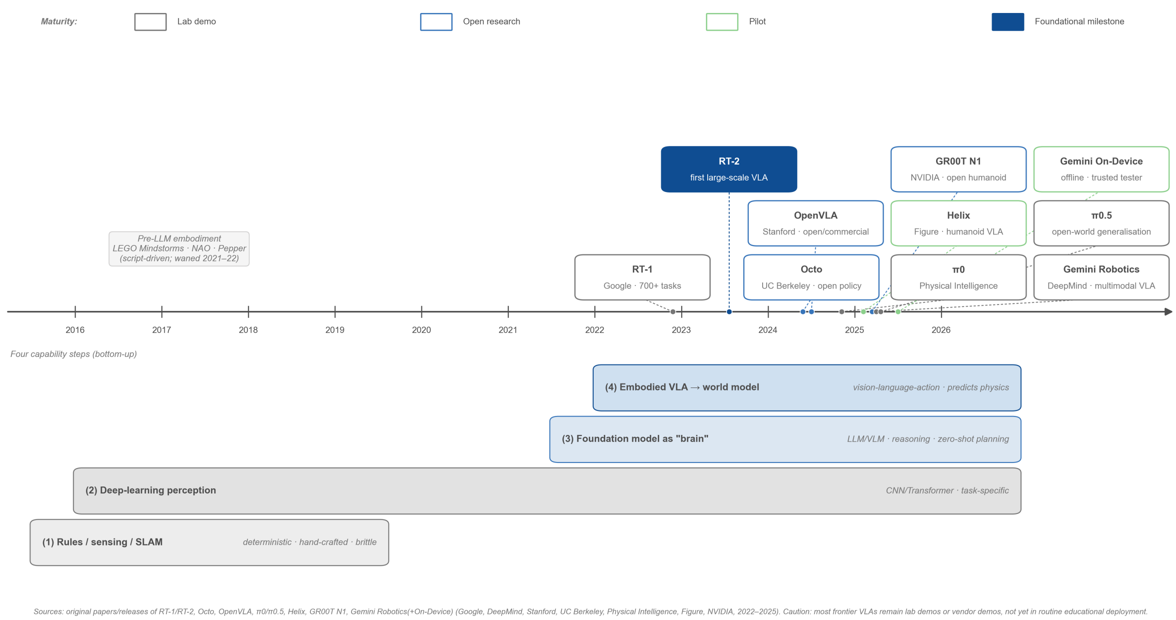


Figure 05.

4.3 Historical-Baseline Bodies: Three Representative Machines of the Pre-Foundation-Model Script Era

To gauge how deep the break runs, look first at what came before it. The three machines below were all landmark products of the pre-foundation-model era, and they share one feature: the hardware body came first, while the cognitive "brain" was driven by rules and scripts.

LEGO Mindstorms / VEX / Thymio — the paradigm of the programmable kit. Centered on block-building plus graphical programming, cultivating computational thinking through rule-based behavior, these were the long-running workhorses of the K-12 stage, with cumulative global deployment in the hundreds of thousands of sets; they are **deployed** and were the largest mature track of the pre-foundation-model era ([51]). One symbolic milestone: LEGO discontinued the Mindstorms line in 2022 — readable as the generational farewell of the "rule-and-script paradigm" in the consumer education market.

NAO — the script-driven small humanoid. Made by Aldebaran/SoftBank, a small humanoid with 25 degrees of freedom (DoF), it works by means of a preset motion library and scripted dialogue, and has seen long-term use in scenarios such as social-skills training for autism spectrum disorder (ASD); it is **deployed** ([52], [53]). But this must be clear: NAO's so-called "intelligence" in the script era was **script-driven, not generative** — it executed branches a human had written, not responses it generated autonomously. Its parent company, Aldebaran, entered bankruptcy liquidation in February 2025 (detailed in Chapter 11), which also signals the commercial-sustainability risk of this generation of social-robot bodies.

Pepper — the marker of a failed business model. SoftBank's 1.2-meter humanoid, once claimed to have emotion-recognition capability and fitted with a touchscreen, saw commercial deployment for a time — roughly 27,000 units cumulatively — but was discontinued in June 2021 (**once deployed, then**

discontinued, [54]). Pepper's exit is the representative event of the frustrated commercial logic of "pre-foundation-model social robots": the hardware body was already in place, but it lacked a sufficiently strong cognitive "brain" to support the open-ended interaction it promised. **The foundation model supplies exactly the link that had been missing** — also the historical starting point for Chapter 6's narrative of "social bodies plugging in an LLM brain."

4.4 A Timeline of Representative VLA / Embodied Foundation Models

The concrete evolution of step four can be presented as a timeline from RT-1 to Gemini Robotics On-Device. This section **mandatorily tags** each system with a maturity label — the key evidence for understanding the "capability–deployment gap."

Model	Institution	Year	Capability highlight	Maturity
RT-1	Google / DeepMind	2022	Transformer maps images and instructions to discrete actions, covering 700+ tasks	demonstrated, lab
RT-2	Google DeepMind	2023-07	First large-scale VLA; action tokenization, chain-of-thought reasoning introduced	demonstrated, lab (foundational milestone)
Octo	UC Berkeley	2024	Open-source generalist policy (27M/93M parameters), diffusion action head	demonstrated, open-source research
OpenVLA (7B)	Stanford et al.	2024-06	First fully open-source, commercially usable VLA; built on 527K Open X-Embodiment trajectories	demonstrated, open-source research (a "model," not a "product")
π0 (Pi-Zero)	Physical Intelligence	2024-10	VLM backbone + flow-matching action expert; 68 tasks across 7 configurations	demonstrated, lab (open weights, controlled evaluation)
π0.5	Physical Intelligence	2025-04	Open-world generalization; cleaning brand-new kitchens/bedrooms "not in the training set"	demonstrated, controlled research (not deployed to users' homes)
Helix	Figure AI	2025-02	Humanoid VLA, dual-system S2 (VLM) + S1 (~200 Hz), ~500 hours of teleoperation data	vendor-claimed / piloted, early (chores from a vendor video, no third-party replication)
GR00T N1/N1.5	NVIDIA	2025-03	"World's first open"	demonstrated /

			humanoid foundation model, dual-system VLA	piloted, open-source research, early ecosystem
Gemini Robotics	Google DeepMind	2025-03	Multimodal VLA across text/image/audio/video, claims origami-level high dexterity	demonstrated, lab
Gemini Robotics On-Device	Google DeepMind	2025-06	Runs on-device, works offline, fine-tunes with 50–100 demonstrations	piloted / demonstrated, restricted trusted-tester program, not a shelf product

A few milestones warrant naming. RT-2 ([4]) is the first large-scale VLA; it tokenized action and introduced chain-of-thought — foundational work, but at a maturity of laboratory demonstration only. The significance of OpenVLA ([5]) is that it is the first fully open-source, commercially usable VLA, trained on the cross-embodiment Open X-Embodiment data — but one must be clear-eyed: it is a "model," not a "deployed product." π0 ([6]) layers a flow-matching action expert atop a VLM backbone, covering multiple configurations and tasks; its upgrade π0.5 ([55]) demonstrated, in controlled research, the open-world generalization of "cleaning a brand-new kitchen that never appeared in the training set" — but this capability **is the product of a controlled research evaluation, not deployment to real users' homes or classrooms**. For the predecessor and open-source comparison, RT-1 and Octo appear in [56] and [57].

On the humanoid side, Figure AI's Helix ([8]) uses a dual system of S2 (VLM slow thinking) plus S1 (~200 Hz fast control) and claims to complete household chores on the basis of roughly 500 hours of teleoperation data; but the "doing chores" demonstration comes from **a vendor release video**, lacks independent third-party replication, and can only be rated vendor-claimed / early pilot. NVIDIA's GR00T N1/N1.5 ([7]), as a "world's first open" humanoid foundation model, has been connected to bodies from Agility, Boston Dynamics, NEURA, and others, and belongs to open-source research and early ecosystem piloting. Google DeepMind's Gemini Robotics ([58]) demonstrated laboratory capability in cross-modal VLA and origami-level dexterous manipulation; its on-device version, Gemini Robotics On-Device ([59]), supports running on the body locally, works offline, and can be fine-tuned with only 50–100 demonstrations, but is currently a restricted program for trusted testers, not a shelf product. The **on-device-inference direction** the latter represents is especially critical for education's privacy protection, offline availability, and low latency; Chapters 5 and 15 continue this thread.

4.5 The Embodied Data Paradigm: Real-Robot Scarcity and "Why World Models Are Needed"

World models address one fundamental constraint, and it must be stated plainly before they are unpacked: **where embodied-policy training data comes from, and how expensive it is**. This constraint determines the availability and price of educational robots directly.

Unlike internet text, "action-paired sensorimotor data" does not exist at internet scale: a robot policy needs synchronized joint angles, gripper forces, camera frames, and task context, which can only be recorded one instance at a time during physical operation — there is no "free web version." The gap in scale is order-of-magnitude: more than 3.9 million industrial robots run worldwide, yet the largest open robot-manipulation dataset holds only about 1 million episodes; the internet data used to train a contemporary large vision–language model amounts to roughly "100,000 years of human experience," whereas the largest single teleoperation robot dataset to date is only on the order of "1 year" ([60], [61]).

To address this "data drought," there are currently three paradigms, each with trade-offs:

- **Teleoperation**: collected by humans one instance at a time, expensive and slow, accounting for less than 1% of total training samples yet carrying most of the weight on "whether the policy works in the real world" — **the most expensive but the most critical** ([62]). - **Simulation**: fills the long tail of rare scenarios; in 2025 research quantified its trade-off — for same-domain manipulation tasks, **roughly 8 simulation samples ≈ 1 teleoperation sample**; but simulation struggles to model wet friction, material deformation, and specular-metal reflections, and contact-rich tasks are where sim-to-real most often fails ([61]). - **Internet and egocentric video**: provides rough action priors and underlying world knowledge, but is only a prior, not a directly executable policy.

The logic of how this chain transmits to education is clear: **the more expensive real-robot data → the more expensive embodied-policy training → the more expensive or weaker the "built-in capability" of education/research robot bodies**. World models and sim-to-real synthetic data are precisely the technical paths hoped to break this cost chain — the context that must be established before unpacking world models in the next section. A conservative boundary must be set: synthetic-data cost reduction remains at the research/early stage, and this blue book lists it as **a trend, not the current state**.

4.6 World Models: Enabling Technology and Research Frontier

The extension of step four is the **world model** — it does not directly control a robot but provides embodied intelligence with "a predictable imagination of the physical world": generating interactive training environments, producing synthetic training data, or serving as a robot's internal predictor for planning, thereby **lowering the barrier to real-robot data collection and accelerating capability iteration**. Between 2025 and 2026 this direction became one of the frontiers most attended to by capital and research across the entire embodied track.

One transversal boundary governs the section: **world models are, without exception, "enabling technology / research frontier," not classroom products**. Their relationship to education is **indirect** —

they drive the maturation of embodied intelligence through cheaper data and faster iteration, and are not a product that can be moved straight into the classroom. Any framing of "world models entering the classroom" is an overclaim this blue book consistently avoids. The term "world model" itself has not yet been classified uniformly — frontier teams such as Fei-Fei Li's World Labs are still writing dedicated pieces to delimit its extension ([63]) — so the discussion proceeds by **three classes of use** rather than blanket assertions about "what world models can already do."

Class one · real-time interactive worlds ("playable worlds / agent training grounds"). These systems generate, from text or images, dynamic environments that can be navigated and interacted with in real time, in a form close to "playable worlds," their main value being to serve as training and evaluation grounds for agents.

- **Google DeepMind Genie 3**: a general-purpose world model that generates, from a text prompt, a dynamic world navigable in real time at 720p / 24 fps, with about 1 minute of "visual memory," and can connect a SIMA agent for long-horizon goal evaluation in the generated environment. The official notes list limitations: a constrained executable action space, unresolved complex multi-agent interaction, an inability to simulate real geography precisely, and continuous interaction currently supported only for "a few minutes." Maturity is **demonstrated / research preview** — its consumer-facing demo opened on a limited basis to AI Ultra subscribers in the United States from January 2026 ([64], [65]). - **Decart Oasis / Mirage / Oasis3**: Oasis is billed as the first real-time, interactive generative video/world model; MirageLSD achieves real-time video-stream transformation with sub-40 ms response; Oasis3 generates photorealistic driving environments in real time and is already an API, initially oriented toward autonomous driving. Maturity is **demonstrated + early commercial API** ([66], [67], [68]). - **Odyssey-1 / Explorer**: pioneering "interactive video," streaming a new frame every 40–50 ms with real-time response to input; Explorer turns any image into an explorable 3D world. The official materials explicitly point to real-simulation uses such as robot training environments. Maturity is **research demo / early preview** ([69], [70]). - **Runway GWM-1 / Luma Ray3**: video-generation vendors are extending toward world models — Runway generates explorable environments in real time while maintaining geometric/lighting/physical consistency, and Luma Ray3 supports start-and-end-frame transitions and character-reference editing. Both sit between "video generation" and "world model" and must be classified with care; maturity is **commercial video-generation product + world-model-direction exploration** ([71], [72]).

Class two · synthetic-data generators ("making training data for robots/self-driving"). These systems do not pursue "playability" but focus on producing physically plausible synthetic training data at scale, directly addressing the cost-chain problem of §4.5.

- **NVIDIA Cosmos** (World Foundation Models): generates "physics-aware" video, supports Text/Image/Video2World, has iterated into the Predict / Transfer / Reason series, and works with Omniverse as a post-training synthetic-data engine for robots and autonomous driving; 1X, Agility, Figure, Uber, and others are among the first adopters. Maturity is **platform released (open models) + developer/enterprise pilot adoption** — currently the world-model platform with the broadest developer ecosystem and the closest to a "usable tool" ([73], [74], [75]). - **Wayve GAIA-2**: a latent-diffusion world model that generates controllable synthetic data for assisted/autonomous driving, able to synthesize rare and high-risk scenarios for stress testing. Maturity is **research release / internal enterprise production use** (Wayve uses it internally for safety testing; [76], [77]).

Class three · self-supervised predictive world models ("the robot's internal planning model"). These systems internalize the world model as the robot's predictor, used for planning and control, and are the class closest to the literal meaning of "cerebellum."

- **Meta V-JEPA 2**: pre-trained on more than 1 million hours of internet video, then adapted with about 62 hours of unlabeled robot video, it uses an internal world model for model-predictive control (MPC) and completes **zero-shot** pick-and-place on a Franka arm in a new lab (from image goals alone, with no task-specific training or reward). It demonstrates the path "video world model → executable robot planning," substantially reducing dependence on real-robot data. Maturity is **research release (open models + benchmark) / lab validation**, not a product ([78], [79]).

Parallel to the three classes is a direction emphasizing **spatial intelligence**: World Labs' (the Fei-Fei Li team's) **Marble** generates exportable 3D environments from text/image/video, released in November 2025 as the first commercial product and closing roughly US$1 billion in funding; it is an **early commercial product / research frontier** ([80]). This blue book lists it alongside world models but classifies it separately as a "spatial intelligence" sub-class to avoid conceptual conflation.

Laying the nine systems out by "class × maturity" yields the map of : real-time interactive worlds mostly remain at research preview / restricted demo, Cosmos among the synthetic-data generators is closest to a usable tool, and the self-supervised-predictive V-JEPA 2 is still at lab validation; step by step, **maturity decreases from the synthetic-data platform toward real-time interactive worlds and then toward end-to-end planning** — the closer to "directly controlling a real robot," the less mature.

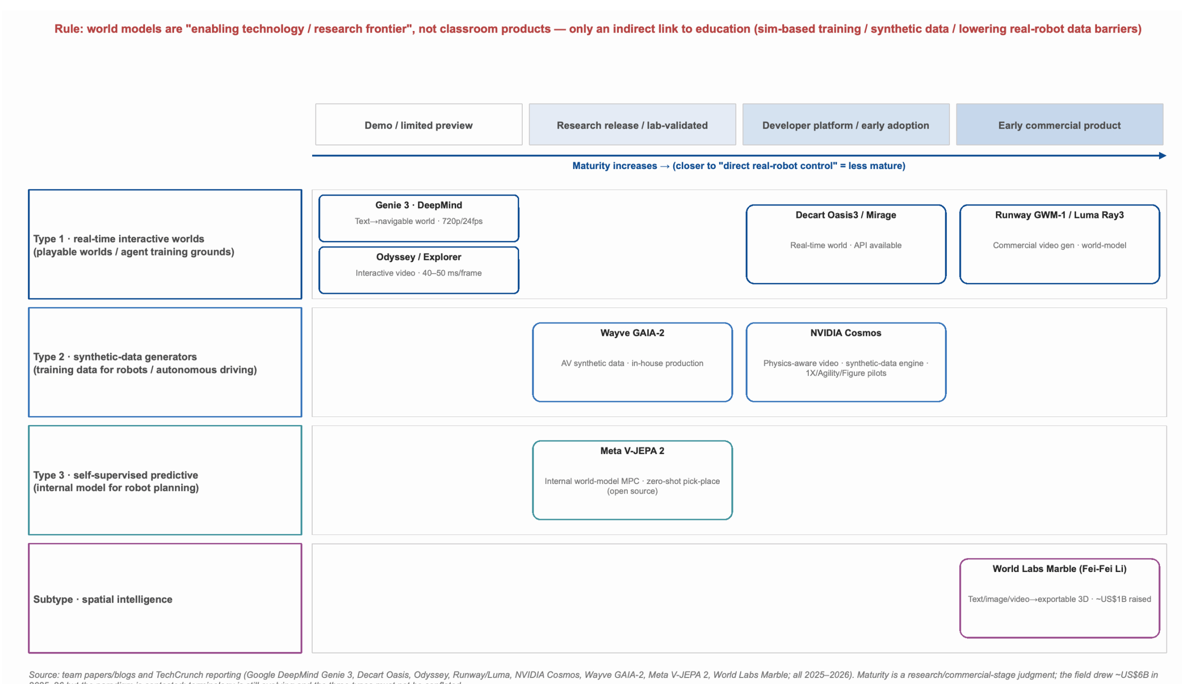


Figure 26.

One cautionary signal belongs in the main text. Between 2025 and 2026 the embodied/world-model direction attracted roughly US$6 billion in investment, yet analysts have already warned that "the scaling paradigm of large language models may not transfer directly to world models" — one cannot take for granted that piling on data and compute will solve embodied intelligence ([81]). There is a gap between capital heat and paradigm certainty, consistent with this blue book's main spine of the "capability–deployment gap": **world models are a promising enabling technology, but whether — and how fast — they can press down the capability and cost curves of embodied intelligence (educational robots included) remains an open question**. For education, the transmission logic that can be written into "future trends / policy recommendations" is this — if world models plus sim-to-real synthetic data mature, they could lower the cost of a single educational robot reaching usable capability and make "doing practical training in simulation" a low-barrier teaching mode (students could practice reinforcement learning / perception / planning without expensive real robots; for related surveys see [82] and [83]); but this is a seed of a trend, by no means the current state.

4.7 An Overview of the Maturity Distribution: The Core Argument for the Capability–Deployment Gap

Merging all the systems of §4.3–§4.6 by maturity, the conclusion is plain at a glance ():

Figure 06.

- **Deployed (commercial supply or independent on-site reporting)**: LEGO Mindstorms / VEX (pre-foundation-model kits, historical workhorses); NAO (the body is deployed, but its "intelligence" was once scripted); **Khanmigo (a software LLM agent, not a robot body)**; Unitree (humanoid bodies for the research/developer market, [84]). - **Industrial piloted**: Figure 02 (a single-customer production line, unrelated to education). - **Piloted / small scale**: Gemini Robotics On-Device; several LLM-plus-social-robot classroom studies (27-student scale, week-long durations). - **Demonstrated / vendor demo / research preview**: RT-1, RT-2, OpenVLA, Octo, π0, π0.5, Helix, GR00T N1, Gemini Robotics, Tesla Optimus, and almost all multimodal learning-analytics systems; **the nine world-model systems also fall, as a whole, within this band** — among them Cosmos is a released platform with developer pilot adoption, V-JEPA 2 and GAIA-2 are research releases with lab/internal-enterprise validation, Genie 3, Oasis, Odyssey, and Runway/Luma are mostly research preview or restricted demo, and Marble is an early commercial product; none constitutes an educational deployment.

From this follows the **core cautious judgment** of this chapter — and of the entire technical part: the most frontier capabilities of 2026 — "foundation model + embodied intelligence + world model" — **remain, for the great majority, at laboratory demonstration and vendor release video, and have not yet entered routine educational deployment**. What has truly entered the classroom and scaled is either the programmable kits of the pre-foundation-model era or the screen-based LLM agent — the largest of which is Khanmigo, whose K-12 student users grew rapidly from about 40,000 in 2023–24 to about 700,000 in 2024–25 (the company projects more than one million in 2025–26), while its total users (including teachers) reached about 1.4 million by April 2025, with roughly 350 US school districts using it daily ([10], [11]). This is the true picture of "**the brain has landed; the embodied body has not**." This blue book must present this "capability–deployment gap" as a basic fact, avoiding treating a vendor demo as an accomplished fact.

4.8 Chapter Summary

This chapter has traced, along four steps, the technical evolution from rule-based scripts to embodied foundation models: the generational break genuinely exists, but it occurs chiefly in the dimension of "capability," and the leap in the dimension of "deployment" is far from synchronously complete. We tagged every frontier system with a maturity label, deepened the three-class distinction of world models (real-time interactive worlds / synthetic-data generators / self-supervised predictive) and their cautious positioning as "enabling technology, not a classroom product," and aggregated in §4.7 the distribution picture of "the great majority of frontier systems remain demonstrated/piloted" — an argumentative anchor to which the second part, and indeed the whole report, repeatedly returns.

Following the foundation-model step, Chapter 5 turns to **agentification**: it analyzes why the first thing deployed at scale on step three was not an embodied body but the software LLM tutor represented by Khanmigo, and discusses the architectural implications of "brain ↔ body separation," real-world samples of hallucination governance, and the responsive direction of on-device inference. Following humanoid VLAs such as Helix/GR00T and the Unitree body, Chapter 6 cautiously characterizes the hardware curves of humanoid and social bodies and the boundary of their educational accessibility. The replacement of the technical-keyword list relative to earlier editions (multimodal large models / embodied intelligence (VLA) / agent orchestration and memory / on-device inference / trustworthiness and alignment) corresponds to the re-definition of Chapter 3 and maps, in Chapter 11, to the value chain's newly added VLA algorithm layer.

Chapter 5 Agentification: When the Software Tutor Reaches Scale Before the Robot Body

> ***The maturity-label convention is the same as in Chapter 4.*** *The chapter's core distinction: the screen-based software LLM agent (already at scale) versus the embodied LLM robot body (still research/pilot) — the two must never be conflated.*

5.1 Introduction: When the "Brain" Lands Before the "Body"

Chapter 4 laid down a key fact while tracing the four technical steps: on the step of "the foundation model as the brain," **the first thing truly deployed at scale was not any robot but a suite of software agents running on a screen.** This runs counter to the popular intuition of a "robot era," yet it is the most important deployment reality in education in 2026.

Agentification is the subject here: the foundation model is no longer a passively responding dialogue box but is organized into an agent with the capabilities of **tool calling, memory, sub-agent collaboration, and role orchestration.** Plug such a "tutor brain" into a robot body and the result is the agentic tutor robot. One boundary holds throughout, though: **the only thing currently deployed at scale is the software agent on the screen, while the embodied tutor-robot body remains almost entirely at research and small-scale pilots.**

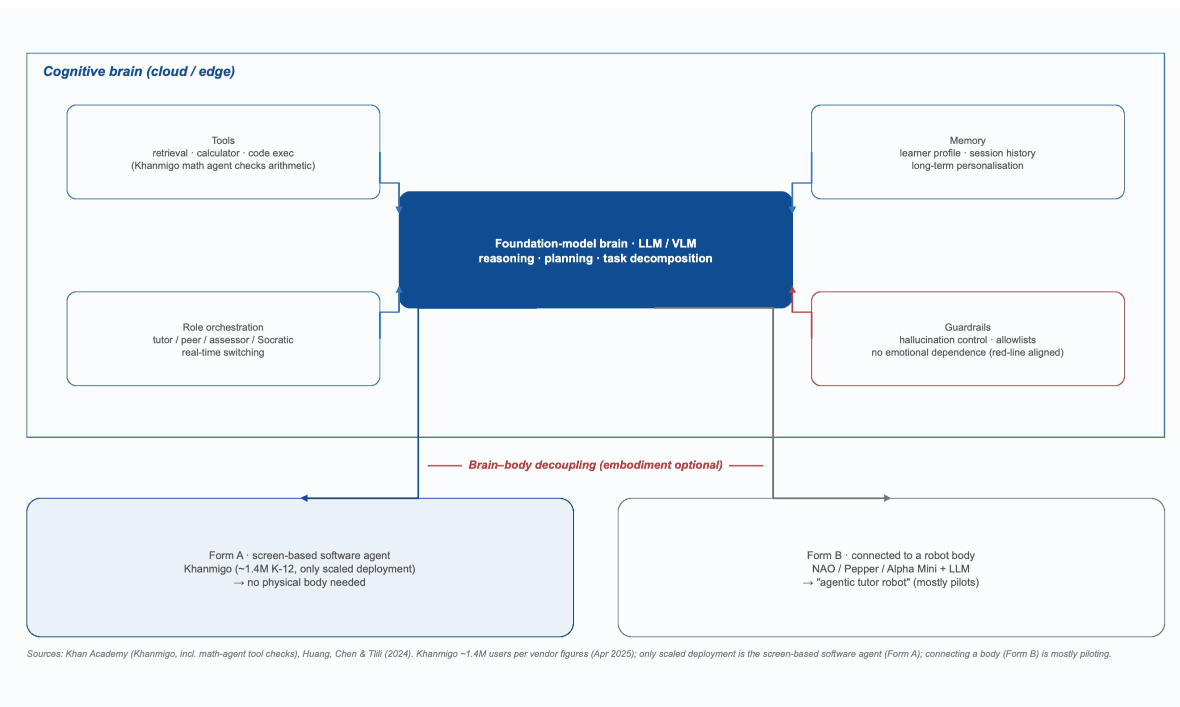


Figure 07.

5.2 The "Tutor Brain" Paradigm: Khanmigo as the Only Sample Deployed at Scale

Khan Academy's Khanmigo is, to date, the **only LLM agent tutor in education deployed at scale,** and the most powerful empirical evidence on the "the brain has landed" side of the capability–deployment-gap narrative ([10]).

Its capability positioning is Socratic guidance: rather than giving answers directly, it advances students through questioning, while batch-generating lesson plans and grading rubrics for teachers. The deployment trajectory is especially telling. Launched as an experimental project in 2023 and scaled from the 2024–25 school year, its **K-12 student users grew rapidly from about 40,000 to about 700,000, while total users (including teachers) reached about 1.4 million by April 2025, with roughly 350 US school districts using it daily;** it has also expanded to India, Brazil, the Philippines, and elsewhere ([11]). Note that the user and district figures are vendor and annual-report figures (mostly registrations rather than precise active counts) and should be cited as such.

One distinction bears repeating: Khanmigo is a **software agent that runs on the screen; it is not a robot body.** Its value lies in showing that the teaching capability of LLM agents has been validated by real school districts in routine settings; only once this "brain" is plugged into a robot body does it constitute an agentic tutor robot. What Khanmigo validates is the "brain," not the "body." Reading its deployment success directly as "educational robots have entered classrooms at scale" is a conflation of scope that this blue book strictly avoids. A further boundary applies: Khanmigo validates **deployment scale, not learning outcomes** — as of the end of 2025 its instructional effect still lacks confirmation by a gold-standard randomized controlled trial (RCT) (the evidence on outcomes is consolidated in Chapter 10), and "scale has landed" does not equal "outcomes are proven." Its state-level uptake is deepening nonetheless: in New Hampshire, roughly 50 districts, about 5,000 teachers, and about 40,000 students are now connected (state-government basis, [85]), a sign that the "tutor brain" is moving from a single platform toward state-level public-education infrastructure.

Khanmigo is not an isolated case but the most representative sample of a **broader phenomenon in which the software agent has reached scale while the body has not.** Contemporary screen-based tutors that sit alongside it include Amira Learning (an AI reading tutor), which on a vendor basis reaches about 4–5 million students across more than 2,000 districts, all 50 US states, and about 19 countries, and which has been rolled out free statewide in Iowa, directly procured for K-5 in North Dakota, and deployed district-wide for K-3 in Newark, with Jordan further incorporating it into a national-level pilot ([86], [87]); and Google's Gemini for Education, which reached about 10 million students and more than 1,000 institutions in 2025 (a vendor reach basis, [88]). Together these samples confirm that **when the "brain" is delivered in pure-software form, scaled deployment is already a fact in 2026** — but none of them is a robot body, and they must still be kept strictly distinct from the embodied tutor robot. To restate the conservative boundary: most of the above user/district/reach figures are vendor or annual-report figures (registrations rather than precise active counts), and the strength of their outcome evidence varies widely (see Chapter 10); "already reaching scale" does not equal "already proven effective."

5.3 A Real-World Sample of Hallucination Governance: Khanmigo's math agent

One of the most criticized risks of LLMs entering education is **hallucination** — the model may confidently produce incorrect content, which is especially dangerous in subjects such as mathematics that have a unique correct answer. Khanmigo offers an engineered response worth recording ([10]).

Its built-in "math agent" uses a mechanism of **tool calling plus a dedicated validation sub-agent:** when a conversation involves a concrete computation, the main model does not output a "mental-arithmetic" result directly but hands the calculation to a deterministic tool or a specialized validation sub-agent, thereby mitigating the LLM's proneness to arithmetic error. The methodological significance of this design is that it shows how to **constrain hallucination through an agent architecture rather than by simply piling on model scale** — replacing the unreliable generation step with reliable tool execution, and doing after-the-fact validation with a sub-agent. This is exactly what distinguishes "agentification" from "a single large-model dialogue": the capability comes from **orchestration,** not from the model's own omniscience.

The math agent significantly mitigates but **does not eliminate** hallucination risk; in open subjects, long-horizon dialogue, and multimodal settings, hallucination governance remains a far-from-closed engineering and governance problem (for the specifics, see the discussion of academic misguidance in Chapter 14).

5.4 LLM-Driven Social-Robot Bodies: A Panorama of Research and Pilots

The exploration of plugging the "tutor brain" into a physical body is already under way, but **all of it remains at the research or small-scale-pilot stage, with none constituting a scaled deployment.** The table below aggregates representative LLM-plus-social-robot-body work in educational settings, tagging each line with its maturity:

System / work	Year	Description	Maturity
Pepper + ChatGPT/GPT-3 teaching new content in a high-school class	2024	Using an LLM to drive Pepper to teach new content in class	piloted / lab study (small sample, [89])
NAO + LLM in special education (gesture / sign language)	2025	An LLM enhancing NAO's gesture and sign-language interaction	lab study ("great potential, not yet fully developed," [90])
Humanoid social robot as a classroom teaching assistant (recognizing engagement)	2025	An LLM-driven humanoid recognizing student engagement and assisting	lab / pilot study ([91])
RoboBuddy (LLM storytelling)	2025	An LLM robot telling stories to children	piloted research (27 students, one-week classroom, [92])
LLM storytelling to spark children's creativity	2024	Investigating the effect of LLM storytelling on creativity	piloted / lab study (peer-reviewed, [93])

Peer-style teachable robot (learning-by-teaching)	2025	Students "teach" a robot to enhance their own retention	lab study ([94])

Three things stand out from this panorama. **The new trend in body choice is to reuse an off-the-shelf low-cost body and add an LLM brain** — researchers mostly adopt mature platforms such as Pepper, NAO, Kebbi, and Furhat rather than developing more complex bodies of their own (see Chapter 6). The logic is economic: as §4.5 noted, the real-robot training data that embodied policies require is scarce and expensive, and developing a complex body in-house and endowing it with end-to-end autonomous manipulation is extremely costly; reusing a mature body and plugging in a language brain sidesteps the most expensive embodied-data-collection step and concentrates innovation at the dialogue and interaction layer. **Current LLM-social-robot research is feasible precisely because it mainly consumes the already-scaled capability of language models, not the expensive capability of embodied manipulation.** Beyond that, **sample sizes and durations are generally very small** — RoboBuddy, with only 27 students over one week, is among the more substantial of them, and most of the work remains in controlled labs or feasibility validation. And **none constitutes a scaled deployment;** together they depict prospect, not outcome. This blue book therefore positions the body of work as **broad in prospect, early in evidence, with no claim of effectiveness whatsoever.** The evidence-based discussion of learning outcomes is consolidated in Chapter 10, and this chapter does not repeat effect sizes.

5.5 Agent Orchestration and Memory: From Fixed Roles to Orchestratable Roles

What agentification brings is not only "plugging into a body" but, more fundamentally, **an upgrade of the interaction paradigm.** Educational robots have long been typecast into three fixed roles — teaching assistant, teaching tool, and learning companion; under the LLM-agent framework, this static classification can be restated as **role orchestration** — a single agent brain can switch identity in real time according to the situation: a Socratic-questioning tutor at this moment, a formative assessor the next, and a learning-data-analytics advisor or a collaborative-learning facilitator the moment after (the full argument and basis for this re-definition appear in Chapter 3, the theoretical anchor being that [9] has already brought chatbots within the scope of educational robots).

The two underlying capabilities that support role orchestration are **memory and orchestration.** Memory lets the agent maintain continuous cognition of the learner's profile, progress, and preferences across sessions — it upgrades one-off question-and-answer into continuous tracking of the same learner, allowing context such as "last lesson you got stuck on finding a common denominator" to carry over and thereby support personalized formative feedback. Orchestration organizes "which tool to call, which role to switch to, and when to hand off to a sub-agent" into a schedulable process: an agent with tool-calling capability can, in the course of a conversation, invoke a calculator in real time, retrieve teaching materials, query a learning-records database, or call a deterministic solver, rather than letting the model "compute from memory." This is exactly the generalization of the math-agent mechanism in §5.3: **the**

educational capability of an agent comes from orchestrating tools, memory, sub-agents, and roles, not from the inherent omniscience of a single model.

This judgment that "orchestration brings capability" already has preliminary support from controlled research. In an exploratory RCT in UK classrooms (165 students aged 13–15, 5 schools, the Eedi platform), Google DeepMind's LearnLM (a Gemini variant fine-tuned for learning settings) had **76.4% of its drafted tutor messages adopted by supervising teachers with zero or only minimal edits,** and students were about 5.5 percentage points more likely to solve a new problem than the control group ([95]). A strict boundary is required: the study had a small sample, was exploratory, and made no claim of significance; 76.4% is a **process-quality proxy indicator, not a learning-attainment effect size,** and LearnLM is a text-based LLM, not a robot — it supports the claim that "an orchestrated and aligned tutor brain can produce teaching content that teachers can adopt directly," not any conclusion about outcomes (outcomes belong to Chapter 10).

"Brain ↔ body separation" is a natural corollary of this architecture — the same tutor brain can serve a screen or a tablet, or drive a social robot; the body is merely one carrier of the brain, not the source of the capability (Fig. 7 characterizes the architecture as "brain / tool calling / memory / role orchestration + brain ↔ body separation"). The separation also explains why the "brain" can reach scale ahead of the "body": a software agent's marginal deployment cost is extremely low — copying a subscription suffices to reach a new student — whereas every additional embodied body means real hardware, maintenance, and embodied-capability-acquisition costs, the last of these being the source of the two common challenges in the next section.

5.6 Two Common Challenges: Hallucination and Compute Overhead

Plugging an LLM brain into an educational body faces two transversal common challenges, both of which belong in the cautious sections.

Challenge one · hallucination. As §5.3 noted, the robot may deviate from the preset and produce incorrect content; in education settings facing children, the cost of incorrect content is especially high (it touches academic misguidance and trust). Khanmigo's math agent offers the mitigation paradigm of "tool calling plus sub-agent validation," but it is far from a general solution; hallucination governance under open subjects and long-horizon interaction remains an open problem.

Challenge two · compute overhead. Large models are hard to run offline on an embedded body — compute, power consumption, and latency all impose real constraints. This challenge is especially acute for education, because the school environment often requires **offline availability, low latency, and data not leaving the premises** (for privacy compliance). The **on-device-inference direction represented by Gemini Robotics On-Device ([59]), mentioned in Chapter 4, is a direct response to this challenge** — running on the body locally, working offline, and fine-tuning with only a few demonstrations. Local small models and on-device inference are therefore listed by this blue book as one of the key breakthroughs for closing the capability–deployment gap in education settings (the trend assessment appears in Chapter 15).

A conservative boundary must be set: on-device inference currently remains at the pilot / restricted-program stage (trusted tester); it points the direction but has not yet pushed embodied tutor robots into scaled deployment.

A further cost constraint, adjacent to "compute overhead" and more fundamental still, must be added: plugging the "brain" into a body is only half the problem; the other half is endowing the body with real embodied-manipulation capability, and the latter is constrained by the **real-robot-data bottleneck** described in §4.5–§4.6. The magnitude of this bottleneck is worth characterizing concretely in this chapter, because it directly determines "why the embodied tutor robot is expensive and hard." Unlike internet text data, the "action-paired sensorimotor data" needed to train an embodied policy (synchronized joint angles, gripper forces, camera frames, and task context) has no ready-made internet-scale counterpart and can only be recorded entry by entry through physical manipulation: although more than 3.9 million industrial robots operate worldwide, **the largest open robot-manipulation dataset holds only about 1 million episodes;** by contrast, the internet data used to train contemporary large multimodal models is roughly equivalent to "100,000 years of human experience," whereas the largest single teleoperated-robot dataset to date is only on the order of "1 year" ([60]). This order-of-magnitude chasm is the data-side root of "the language brain has reached scale while the embodied body has not."

The cost-reduction chain is composed of trade-offs among three data paradigms ([61]): **teleoperation,** collected manually entry by entry, is the most expensive and slowest yet carries most of the weight of "whether a policy works in the real world"; **simulation synthesis** is used to fill the long tail of rare scenarios, and in 2025 research quantified a conversion of "about 8 simulated samples ≈ 1 teleoperated sample" (for in-domain manipulation tasks), though it still fails most readily in contact-rich settings such as wet friction, deformation, and specular reflection; and **motion capture and first-person video** provide action priors, but only priors, not executable policies. The indirect value of world models (detailed in Chapters 4 and 15) lies precisely here — generating diverse embodied trajectories at scale through a fully synthetic pipeline, bypassing manual teleoperation, thereby lowering the cost for "a single educational/research robot to reach usable capability" and making "doing practical training in simulation" a low-barrier teaching method (students can practice reinforcement learning, perception, and planning without an expensive real robot) ([75], [82]).

The two levers can thus be set side by side: on-device inference lowers the brain's **deployment cost,** while world models and sim-to-real synthetic data, if they mature, hold promise for lowering the body's **capability-acquisition cost** — the two are viewed by this blue book, from the "brain accessibility" and "body accessibility" ends respectively, as potential levers for closing the capability–deployment gap. But, again, a conservative boundary: both currently remain at the pilot / research stage — a trend, not the current state (the trend assessment appears in Chapter 15); and one must be especially wary of describing world models as "an education product already in the classroom" — they are an **enabling technology** whose relationship to education is indirect (lowering the real-robot-data barrier, synthesizing training data, supporting simulation-based practical training), and in no sense a classroom product.

5.7 Chapter Summary

This chapter has established the core picture of agentification: through tool calling, memory, sub-agent validation, and role orchestration, the foundation model is upgraded into a "tutor brain." This paradigm **first reached scale in the form of a software agent** — Khanmigo, with about 1.4 million total users (including teachers) and daily use in roughly 350 US school districts, is the only sample at scale, but it is screen-based software rather than a robot body, and its outcomes have not yet been confirmed by a gold-standard RCT. The exploration of plugging the same "brain" into a physical body, though already under way, remains entirely at research and week-long pilots — **broad in prospect, early in evidence.** Hallucination and compute are the two common challenges along this path; on-device inference and world-model synthetic data are, from the brain and body ends respectively, the responsive directions worth tracking.

Following the architectural implication of "brain ↔ body separation," Chapter 6 turns to the "body" side, cautiously characterizing the hardware curves, cost inflection points, and educational-accessibility boundaries of humanoid and social bodies — and making clear that there is a non-negligible distance between a low-cost social body that can plug in an LLM brain and the "concept/demonstration" of industrial-grade humanoids entering the classroom. The LLM teaching evidence touched on in this chapter (Tutor CoPilot, LearnLM, and others — all text-based LLMs, not robots) is handled uniformly in Chapter 10; the governance specifics of hallucination and academic misguidance belong to Chapter 14.

Chapter 6 The Full Spectrum of Embodied Bodies: Humanoids, Dexterous Hands, and Quadrupeds

> ***Scope statement for this chapter****: this chapter covers only* ***embodied hardware*** *— the degrees of freedom, cost, parameter curves, and educational-accessibility boundaries of three classes of body: humanoids, dexterous hands, and quadrupeds. The market size and shipment forecasts of these bodies belong to Chapter 12; vendor financials and competitive rankings to Chapter 13; specific vendors' educational-deployment cases to Chapter 9; and the value-chain restructuring of upstream "bottleneck" components to Chapter 11 — this chapter only points to those as background, without re-elaborating them, to avoid cross-chapter repetition. > >* ***General rule on cooling the humanoid-education narrative (running through this chapter)****: whole-machine specs, prices, and shipments are all tagged "vendor/media figures as of 2025–2026, not independently verified";* ***claims of the form "humanoid-robot teachers entering classrooms at scale" are strictly prohibited****. The humanoids that can enter education (Unitree G1, Fourier GR, Booster T1, and others) are chiefly* ***teaching/research platforms, targets for teacher training, and competition platforms*** *for universities and research; roughly three-quarters of Unitree's humanoid revenue comes from research-and-education customers; large-scale K-12 use still relies on small programming/social bodies. The iron rule on currency: Chinese financials' "亿元 (100 million yuan)" ≠ billion — RMB 100 million ≈ US$14 million.*

6.1 Introduction: When the "Brain Problem" Is Partly Solved, the Focus Returns to the "Body Side"

Chapter 4 noted that Pepper's discontinuation marked the frustration of the "pre-foundation-model social robot" business model — the hardware body came first but lacked a sufficiently strong cognitive brain; its parent company, Aldebaran, entered bankruptcy liquidation in February 2025 (cumulative debt of about €150 million), and the intellectual property and assets of NAO/Pepper were acquired at auction by Shenzhen RealMan Microvision, writing a belated footnote to the "body-first" route ([54]). Chapter 5 then established that "brain ↔ body separation" is a natural corollary of agentification, and that the body is merely one carrier of the brain.

These two threads meet here: once foundation models partly solve the brain problem, the **hardware curve, cost structure, and educational accessibility of the body side** become the new factor deciding whether embodied educational robots can land. A cautious main line governs the chapter from the outset — hardware is indeed maturing fast and cost is indeed falling, pulled down by Chinese supply chains, yet **none of this amounts to "humanoid robots will enter primary and secondary classrooms at**

scale." In 2025 — the year widely called the humanoid "year of mass production" — humanoid mass production points almost entirely toward industry, logistics, and retail, while humanoid landing in education classrooms is near zero. The embodied platform genuinely entering education "in the present tense" is instead the cheaper, more motion-control-mature **quadruped robot (robot dog)**.

Sources: public vendor specs/prices, Unitree G1 EDU announcement, SoftBank Robotics Pepper end-of-life notice (2021), industry humanoid cost breakdowns (2026). Price/DoF/capability are vendor or secondary figures, not independently verified; capability is qualitative.

Figure 08.

This chapter splits embodied bodies into three classes for examination one by one — **humanoid (full body), dexterous hand (the manipulation end-effector), and quadruped (the mobile platform)** — strictly distinguishing the four maturity tiers of "deployed / piloted / demonstrated / vendor-claimed" within each class, and finally lands on a gap chart of "mass production vs educational landing."

6.2 The Evolution of Social Bodies: From NAO/Pepper to "a Low-Cost Body + an LLM Brain"

The evolutionary path of social bodies completes the picture before the full spectrum of the three classes, and corroborates the judgment that "the body is merely a carrier."

The first-generation representatives are NAO and Pepper: complex bodies developed in-house, dependent on script-driven operation. Pepper was discontinued in 2021 (about 27,000 units cumulatively, [54]), while NAO still sees use thanks to its long accumulation in scenarios such as ASD — a historical cumulative of about 20,000 NAO and 17,000 Pepper units sold to about 70 countries, with more than 17,000 cumulative units into the education market. Their common lesson is this: **simply making the body more complex and more "human-like" cannot sustain the commercial value of open-ended**

interaction; and Aldebaran's liquidation further exposed the sustainability risk of the "heavy-asset hardware body + weak brain" model (who takes over the operation and service of the existing fleet has itself become a problem).

Hence the new trend of 2026 — **keep the low-cost social body and shift investment toward plugging in a cloud or on-premises LLM brain** rather than developing a more complex body. Research and product practice now more often take off-the-shelf platforms such as Kebbi, Minibo, and NAO and layer multimodal LLM capability on top (for the full panorama of "LLM + social robot" research, see §5.4). The industrial implication is that the body link trends toward standardization and low cost while the center of value migrates toward the "brain (model/agent)" and the "service (curriculum/operation)" — a value-migration thread developed further in Chapter 11 on value-chain restructuring.

6.3 The Full Humanoid Spectrum: 20-plus Overseas and Chinese Models

In 2025–2026, whole-machine humanoids evolved from single star products into a comparable full spectrum. The text below organizes more than 20 representative models by the specs, price, shipments, and maturity of two camps, overseas and Chinese. Three reading conventions apply throughout. The whole-machine parameters mostly come from vendor specs and secondhand aggregation, and prices and shipments are rapidly changing data, all tagged "vendor/media figures as of 2025–2026, not independently verified," with the emphasis on trend and order of magnitude rather than exact quotes ([96], [97]). "Mass production/deployment" points almost entirely toward industry, logistics, and retail, not the education classroom. And the humanoids that do enter education are, for the great majority, university/research teaching-and-research platforms, not "robot teachers."

The overseas camp is characterized by the coexistence of "industrial deployment + demonstration claims":

- **Tesla Optimus (Gen 3)**: the hand is claimed to have 22+ DoF, the whole machine about 50 actuators, with a hand technology route of "planetary gearbox + roller screw + tendon drive" hybrid; the V3 formal mass-production target is set for summer 2026, with a capacity target of 1 million units per year and a volume-production cost target of US$20,000–25,000 per unit (current per-unit manufacturing cost estimated at US$50,000–100,000). Maturity: **demonstrated + vendor-claimed** — the mass-production price and capacity are both vendor targets, with no product in stock ([98]). - **Figure 02 / 03 (Helix VLA)**: the 02 has a 16-DoF single hand, payload of about 20 kg, and battery life of about 5 hours; it already supports production on the BMW Spartanburg line, having helped build 30,000+ X3 vehicles (10 hours per shift, moving 90,000+ sheet-metal parts). Maturity: **industrial pilot / early deployment** (an automotive line, not education and not the home); Figure 03's household-chore capability is chiefly a vendor release demonstration ([99]). - **1X NEO**: only 29.94 kg, two 22-DoF hands, an early-bird outright price of US$20,000 or a US$499/month subscription, launching first in the United States in 2026. It must be tagged as **a cautious case**: its complex tasks (such as folding laundry) depend 100% on "Expert Mode" remote teleoperation, with only "opening doors / tidying cups and dishes" tagged as autonomous — the so-called "autonomy" is in fact teleoperation, and one supervisor manages eight remote operators who can see in-home video used for training, which has already prompted privacy controversy. Maturity: **demonstrated + pre-sale (claimed autonomous, in fact teleoperated)** ([100], [101]). - **Apptronik Apollo**: 1.73 m / about 73 kg, payload of about 25 kg, Jetson AGX Orin (275+ TOPS), connected to NVIDIA's GR00T foundation model; it runs logistics pilots at Mercedes-Benz, GXO Logistics, and Jabil; it raised US$520 million in 2025 at a US$5.5 billion valuation. Maturity: **industrial pilot** (logistics, not education) ([102]). - **Boston Dynamics electric Atlas**: the fifth-generation all-electric, with 56 DoF, a 50 kg payload, a 2.3 m arm span, and autonomous battery swapping; all of its 2026 capacity is already committed to Hyundai and Google DeepMind, opening to external customers only in 2027, with a per-unit price estimated by the industry at US$150,000–420,000. Maturity: **first commercial deployment (industrial, with committed customers)** ([103]). - **Agility Digit**: v4 is 175 cm / 60 kg with a payload of about 16 kg, and v5 raises the payload to about 23 kg; it is in commercial deployment at Schaeffler, GXO, Toyota, Amazon, and others, with 65,000+ cumulative operating hours across nine customer facilities and 100,000+ totes moved in a single deployment. Maturity: **commercial deployment (logistics, "beyond pilot")** — the most substantially deployed of the overseas humanoids ([104]). - **Sanctuary Phoenix**: a 21-DoF hydraulically driven hand including a 7-unit micro-pneumatic tactile array (able to sense 5 millinewtons). Maturity: **demonstrated** (an in-hand-manipulation milestone; the hydraulic route is a minority approach) ([105]).

The Chinese camp is characterized by "supply of research/education platforms + the start of industrial mass production," with prices generally below those overseas:

- **Unitree (G1 / H1 / H2 / R1)**: the G1 is about 35 kg / 23–43 DoF, starting at RMB 99,000, with the EDU version including a Python/C++/ROS 2 SDK + Jetson Orin (quoted at about US$43,900–73,900); the R1 (2025-07) has 24–26 DoF, starting at RMB 39,900; the H2 (2025-10) is 1.8 m / 70 kg / 31+ DoF. In 2025 it shipped 5,500+ units to end users, ranking first worldwide (a share of about 32.4%). Maturity: **already in commercial supply (the workhorse for research/education/development)** ([106], [107]). - **Agibot (Yuanzheng A2 / Lingxi X2)**: the A2 is 1.69 m / 40+ DoF, with a youth edition at RMB 168,000–198,000; the X2 is about 1.3 m / 25–31 DoF, with a youth edition at RMB 98,000; the 5,000th mass-production unit rolled off the line in December 2025. Maturity: **already in mass production (chiefly industrial/commercial)**; its "hundred cities, ten thousand schools" education program remains a claim/early stage ([108]). - **Fourier (GR-2 / GR-3)**: the GR-2 is 175 cm / 63 kg / 53 DoF (including a 12-DoF tactile dexterous hand), estimated at ≥ US$150,000 (B2B only); the GR-3 (the "Care-Bot") has ≤ 55 DoF and is priced at > RMB 200,000 (about US$27,500), turning toward emotional-interaction scenarios (including education). Maturity: **already deployed (research/rehabilitation/development platform)** ([109]). - **UBTech Walker S2**: an industrial humanoid, the world's first with autonomous battery swapping (3 minutes); mass-production delivery began in November 2025, with a target of 500 units within the year and a 2026 capacity of 5,000 units, and cumulative orders for the Walker series exceeding RMB 800 million (about US$112 million; **"亿元" is not billion**). Maturity: **mass-production delivery begun (industrial)**; UBTech's education line still continues small programming bodies such as Alpha Mini and Wukong into primary and secondary schools ([110]). - **Robot Era (STAR1)**: a payload of 160 kg, 55 active DoF, a running speed of 3.6 m/s; a Series A of nearly RMB 500 million, a new round of > US$200 million in April 2026, and 200+ units delivered as of June 2025, with 50%+ of orders from overseas. Maturity: **already delivered (chiefly research/development) + vendor-claimed** (valuation of about RMB 10 billion) ([111]). **- EngineAI (PM01 / SE01)**: the PM01 is 1.38 m / 24 DoF at US$12,000 (RMB 88,000); the SE01 is 1.7 m / 32 DoF at US$20,000–30,000; an education edition is explicitly offered. Maturity: **already in supply (research/education/commercial, open-source platform)** ([112]). **- Booster (T1 / K1)**: the T1 is 118 cm / 23 DoF (41 DoF with the dexterous hand), Jetson AGX Orin, about US$34,000; the K1 is about US$12,500; the RoboCup 2025 champion platform, **in use at 70+ universities/research institutions**. Maturity: **already in supply (the workhorse for research/education/competition)** ([113]). **- Galbot (G1)**: a wheeled humanoid, 173 cm, 10-hour battery life, 5 kg payload; landed in smart retail and unmanned pharmacies (RMB 700,000 per unit on the pharmacy side), with an embodied-intelligence joint lab (the EPIC Lab) co-built with Peking University. Maturity: **piloted / early deployment (retail/pharmacy) + a deep industry–academia–research platform** ([114]). **- Songyan Power (N2 / E1)**: bipedal humanoids with a clear education-and-research positioning; within one month of the Beijing humanoid marathon, intent orders exceeded 2,000 units with a contract value of over RMB 100 million, **with the largest customer group from the education sector** (the orders are intent, to be treated cautiously). Maturity: **piloted / early delivery (education-and-research positioning)** ([115]).

Aggregating these 20-plus models into a matrix () yields three judgments. **DoF has generally risen to 20–56**, with electric drive (frameless torque motors plus harmonic/planetary reducers) now mainstream and a few overseas players such as Sanctuary taking the hydraulic route. **Overseas humanoids enter almost only industry, while Chinese humanoids mostly supply in the form of research/education platforms**, at prices generally only a fraction of overseas ones (Unitree R1 at RMB 39,900 against the Atlas at the US$150,000 level). And most crucially, **"mass production" refers almost entirely to industry, logistics, retail, and research, and not a single one is a "humanoid teacher entering primary and secondary classrooms at scale."** Optimus's consumer price, Figure 03's household-chore capability, and 1X NEO's "autonomy" all remain vendor figures or video demonstrations (NEO's "autonomy" is in fact teleoperation), at some distance from "purchasable, deployable, and third-party-reproducible."

Figure 23.

6.4 The Dexterous-Hand Section: The Manipulation Bottleneck, Tactile Density, and China's Price Dimensional Reduction

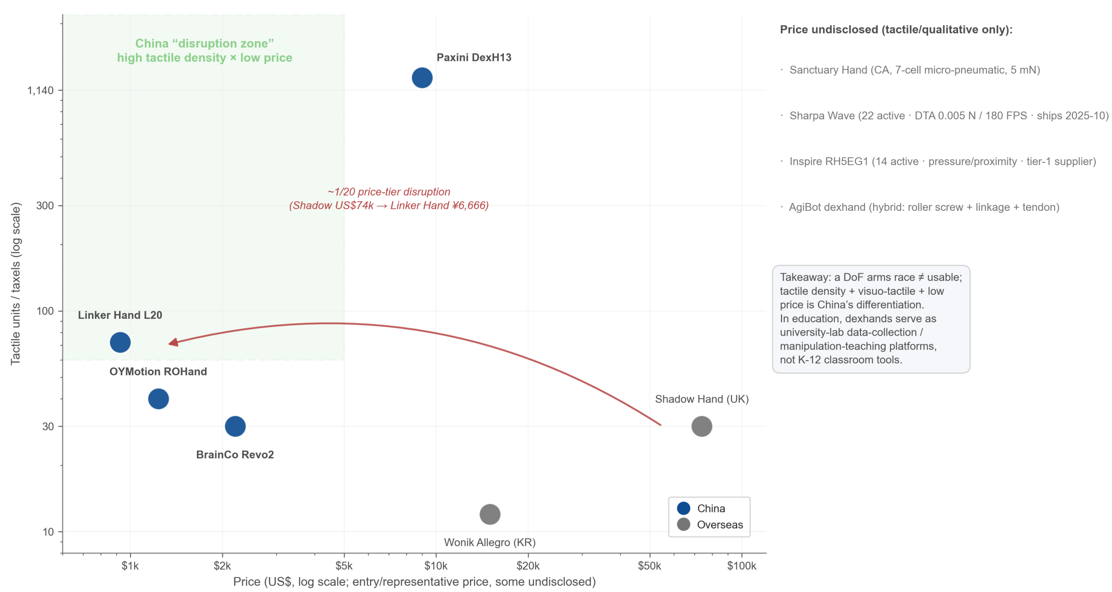
If the motion control of whole-machine humanoids is being engineered fast, then the **dexterous hand remains the real bottleneck of a humanoid's ability to "do things"** — the success or failure of grasping, assembly, and household chores all rests on this hand. The dexterous hand is also a link that warrants its own section, because it is **one of the few fields in which Chinese vendors run alongside, or even ahead of, the West**.

The dexterous hand has three technology routes: **direct drive (linkage/gear, rigid, easy to manufacture and maintain), tendon drive (cable-driven, compliant, human-like but hard to maintain), and hybrid (linkage + tendon, screw + tendon)**. Tactile sensing has become a mainstream standard fitting, and visuo-tactile sensing is regarded by the academic community as the most promising branch. The ten players below are compared:

Four overseas: Shadow Hand (UK, 20 actuated + 4 underactuated, tendon drive, Hall-effect tactile taxels, about US$74,000–100,000, already commercialized for high-end research, [116]); Wonik Allegro (South Korea, 16-DoF direct drive, omnidirectional fingertip pressure sensing, starting at US$15,000, the workhorse for research/education, [117]); Sanctuary Hand (Canada, 21-DoF hydraulic + 7-unit micro-pneumatic tactile, demonstrated, [105]); Sharpa Wave (22 active DoF, a proprietary dynamic tactile array (DTA) with 0.005 N pressure sensing / 180 FPS / spatial resolution < 1 mm, entering mass production and shipping in October 2025, [118]).

Six Chinese: Inspire-Robots (RH5EG1, 14 active DoF, servo-cylinder linear drive + pressure/proximity sensing, a leading supplier of humanoid end-effectors, already in mass production, [119]); BrainCo Revo2 (11 DoF / 6 active, only 383 g, 20 kg payload, built-in 3D tactile sensing of hardness/texture/orientation/proximity, an annual capacity of 30,000 units, priced at a claimed one-fifth to one-seventh of comparable international products, mass-production-ready, [120]); LinkerBot L20 (20 DoF / 16 active, linkage drive, **a 72-unit 200 FPS tactile array** plus optional electronic skin, **starting at RMB 6,666**, with 10,000+ units delivered in 2025, [121]); PaXini DexH13 (16 DoF / 13 active, coreless motors, the highest tactile density in the industry — **1,140 GEN2 ITPU tactile units and 3,420 tactile signal channels**, already commercialized, [122]); OYMotion ROHand (6 active DoF, a high-density tactile array of 0.1–25 N, **starting at RMB 8,888**, a dual line of value-for-money plus prosthetics, already on sale, [123]); Agibot's dexterous hand (hybrid drive: roller screw + linkage + worm gear + tendon, a representative of the hybrid route, [124]).

Plotting "tactile density × price" on the same positioning chart (), the position of Chinese vendors is striking: **PaXini with 1,140 tactile units, LinkerBot with 72 units at 200 FPS, and BrainCo with 3D tactile sensing already lead on tactile density; and on price, LinkerBot at RMB 6,666 and OYMotion starting at RMB 8,888 form a dimensional reduction of nearly an order of magnitude of 1/20 against Shadow's roughly US$74,000 (about RMB 530,000)**. An IROS 2025 industry roundup even ran under the headline "what Tesla has not yet done, Chinese vendors have" (**a media account containing promotional elements, to be cited with caution**, [124]).



Source: vendor sites/specs and IROS 2025 media round-ups (Shadow Robotics, Wonik, Sanctuary AI, Sharpa, Inspire, BrainCo, LinkerBot, Paxini, OYMotion, AgiBot). Prices are entry/representative (RMB converted at ~7.7/US$; undisclosed listed at right); taxel counts are vendor figures; IROS media round-ups include promotional claims — read with caution.

Figure 24.

Two sober judgments follow. **A DoF arms race ≠ usability.** High DoF (20–22) appears mostly in demonstrations and research, whereas industrial production lines value payload and reliability more — Tesla Optimus Gen 3's choice of a "screw + tendon" hybrid over a purely high-DoF scheme reflects exactly that production-line orientation; compete on DoF, but compete even harder on usability. **And in**

education and research the dexterous hand is a platform, not a teaching aid. It enters university laboratories chiefly as an **embodied-intelligence data-collection and manipulation-teaching platform** (teleoperation plus imitation learning); Allegro and Shadow are the classic research platforms, and low-priced Chinese dexterous hands are pulling dexterous-manipulation research into more university laboratories. **It is not a primary/secondary-classroom teaching aid**, where the mainstay remains small programming and social bodies.

The price curve of the dexterous hand is ultimately determined by upstream components — **the planetary roller screw (the most expensive and hardest-to-localize link of a humanoid, accounting for about 28.6%–35% of cost), the harmonic drive (still dominated in share by Japan's Harmonic Drive, with a localization rate of about 30% → a planned 40%), the frameless torque motor, and the six-axis force/torque sensor (where localization leads, the six-axis force localization rate having risen from 19% to 57.8%)** together constitute the cost structure ([125], [126]). The rapid localization of the Chinese supply chain is the underlying fulcrum for lowering dexterous-hand cost and thereby improving educational/research accessibility; but the screw and high-end harmonic drive remain bottlenecks that determine the final price of "affordable for a university laboratory." **The full value-chain panorama of the upstream bottlenecks belongs to Chapter 11; this chapter offers only a one-sentence pointer, without re-elaboration.**

6.5 The Quadruped Section: The Most Realistic Embodied Platform for Entering Education Today

Among the three classes of body, the one genuinely entering education "in the present tense" is neither the humanoid nor the dexterous hand but the **quadruped robot (robot dog)**, for three reasons: it is cheaper (educational kits at the thousands-of-dollars level against humanoids at the tens-of-thousands level), its motion control is basically engineered (already shipping at scale in power-grid inspection and emergency rescue), and it **has actually entered Chinese vocational and university classrooms for perception and algorithm practical training**. This makes it one of the **few closure points** in the "capability–deployment gap" of Chapter 4.

Quadrupeds can be divided into three classes — **consumer/education grade, research/teaching mid-weight, and industrial/defense grade** — and the six players below cover the full spectrum, with the education-edition price highlighted:

- **Unitree Go2 (Air/Pro/EDU/EDU Plus)**: the consumer + education workhorse. **The Go2 Air starts at about US$1,600 and the Pro at about US$2,800; the Go2 EDU standard edition is about US$5,990** (including an Orin Nano at 40 TOPS, a full Python/C++/ROS 2 SDK, a RealSense D435i, and foot-end force sensors); the EDU Plus is about US$16,000–22,500 (upgraded to an Orin NX at 100 TOPS, with an optional 6-DoF arm). Maturity: **mass-produced / already commercialized (consumer + research-and-education)** ([127], [128], [129]). - **Unitree B2/B2-W**: the fastest industrial-grade quadruped (about 6 m/s), payload of about 40 kg (the B2-W wheeled-foot version is IP67 with a payload up to 120 kg); industrial/research mass production. - **DEEP Robotics Jueying Lite3**: a research/teaching mid-weight, with an open SDK +

ROS 2 interface (motion/perception/reinforcement learning), and an **education-and-research edition starting at about US$2,890** — one of the lowest-barrier research-grade platforms for entering education today. Maturity: **mass-produced (education/research)** ([130]). - **DEEP Robotics Jueying X30**: the industrial flagship, about 56 kg, IP67, with an inspection range of ≥ 10 km; already deployed at scale in the Wenzhou underground power-grid corridor and, as **the first Chinese quadruped to go overseas into a foreign power system** (the underground transmission network of Singapore's SP Group, nicknamed "SPock," saving 480+ hours of manual inspection per year). Maturity: **deployed (industrial inspection, in commercial delivery)** ([131], [132]). - **Weilan BabyAlpha A2**: a consumer/home-companionship quadruped with a built-in large language model that can answer children's questions and switch between Chinese and English, with the base edition starting at RMB 6,099–8,099. It must be strictly distinguished: this is **a "learning-companion product" for the home-consumer setting, which does not equal a school educational deployment**. Maturity: **mass-produced (consumer/home, not a classroom teaching product)** ([133]). - **Overseas comparison**: Boston Dynamics Spot (the Explorer Kit starting at about US$74,500, a full configuration at about US$150,000–195,000, with 480+ universities/research institutions worldwide holding a Spot Academic license as of Q4 2025, [134], [135]); ANYbotics ANYmal (starting at about US$75,000, industrial inspection, [136]); Ghost Robotics Vision 60 (about US$225,000–235,000 per unit, chiefly defense and unrelated to education, used only as a price-ceiling anchor, [137]).

Laying these products out by "educational-landing status × price tier" (), the educational accessibility of the quadruped relative to the humanoid is plain: **the Go2 Air at about US$1,600, the Lite3 at about US$2,890, and the Go2 EDU at about US$5,990 are far below the humanoid level of tens of thousands of dollars, let alone Unitree G1 EDU's US$43,900–73,900**. This price tier, together with the open SDK + ROS 2 + point-cloud/visual data streams, makes the quadruped naturally suited to machine-learning/perception/reinforcement-learning practical training.

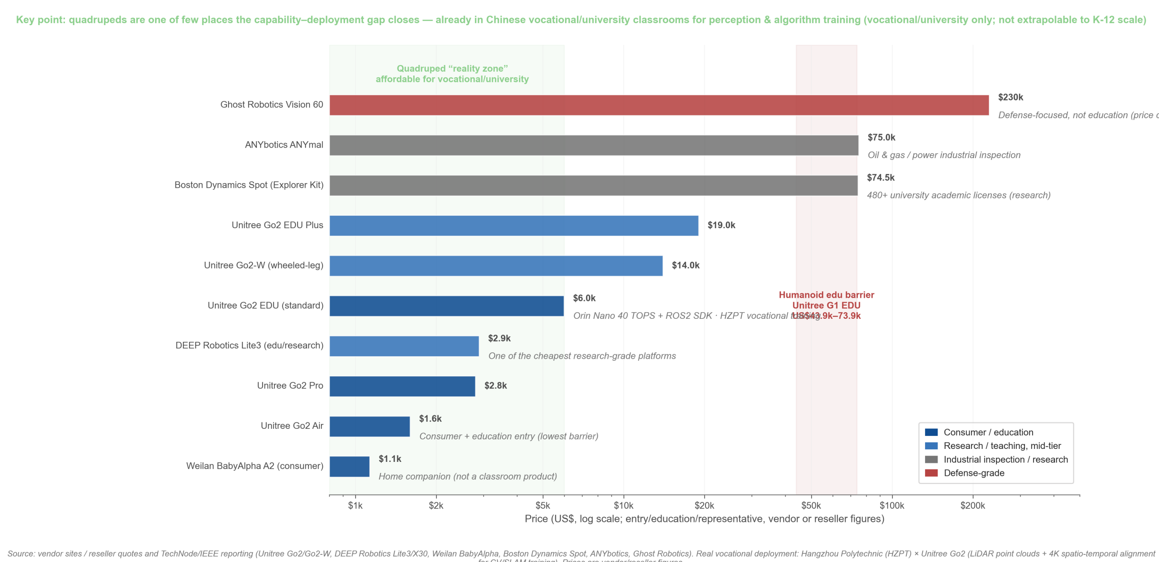


Figure 25.

More crucially, there is **real landing evidence**, the decisive difference between the quadruped and the humanoid:

- **Hangzhou Polytechnic (HZPT) × Unitree Go2**: the School of Internet of Things Technology introduced the Go2 into professional teaching, with students using the **lidar point cloud + 4K camera footage relayed by the robot dog to perform spatiotemporal alignment and write algorithms to identify targets of campus rule-violation behavior**; teachers said it lets students "grasp machine learning and computer vision in a more intuitive way," turning abstract concepts into hands-on engineering challenges. Maturity: **deployed (vocational-college classroom practical training)** — this is a **genuine closure point** of the capability–deployment gap ([138]). **- RoboUniversity**: an educational curriculum and certification system from Unitree in collaboration with a Stanford professor and OpenMind, covering programming fundamentals through advanced robotics integration, oriented to the Go2/G1 platforms ([127]; deployment-case details in Chapter 9). **- iFLYTEK Smart Education's "Future Classroom"**: a quadruped robot dog + a virtual digital human entering a campus science-and-technology festival for display. Maturity: **piloted / display** (a science festival / science popularization, not routine classroom teaching) ([139]).

The educational-landing maturity of quadrupeds versus humanoids can be distilled into the table below:

Dimension	**Quadruped (robot dog)**	**Humanoid**
Motion-control maturity	Basically engineered and solved, deployed at industrial scale	Still iterating fast, mostly pilots
Educational price barrier	Low: Go2 Air about US$1,600, Lite3 about US$2,890, Go2 EDU about US$5,990	High: Unitree G1 EDU about US$43,900–73,900
Educational-landing status	Already in vocational/university classrooms (HZPT practical training; RoboUniversity certification)	Chiefly a university research-and-teaching platform (about three-quarters of revenue from research-and-education); large-scale K-12 use remains small programming robots
Maturity verdict	Mass-produced + deployed (education in the present tense)	Deployed (industrial pilot) + university research platform (educational scale-up still conceptual)

Market figures can serve as side evidence: China's robot-dog sales in the first three quarters of 2024 were about 23,300 units, up 72.22% year over year, with a forecast of about 394,800 units by 2031 (**a research-firm forecast figure, to be tagged**); and within the educational-robot industry, non-humanoid (including quadruped) forms dominate thanks to affordable prices and STEM fit (per Grand View's definition, [140], [141]).

But a bottom line on usage must be held, consistent with the humanoid red line: one may write that "**the quadruped is the most realistic platform for embodied intelligence entering education today, already used for perception and algorithm practical training in Chinese vocational colleges/universities**"; one **must not write** that "robot dogs have entered primary and secondary classrooms at scale" — they are currently confined to vocational/university settings, while primary and secondary schools still rely on small programming robots, and this cannot be extrapolated. Weilan's

BabyAlpha "learning companionship" is home consumption, and still less can it be conflated with a school teaching deployment.

6.6 The Year of Mass Production vs. the Education Classroom: A Summary of the Body Gap

Aggregating the hardware evidence on the three classes of body — humanoid, dexterous hand, and quadruped — and overlaying the gauge of "maturity," yields the chapter's core judgment: **2025 is the humanoid's "year of mass production," yet humanoid landing in the education classroom is near zero; there is a clear gap between the rising curve of body capability and the curve of bodies entering the education classroom** ().

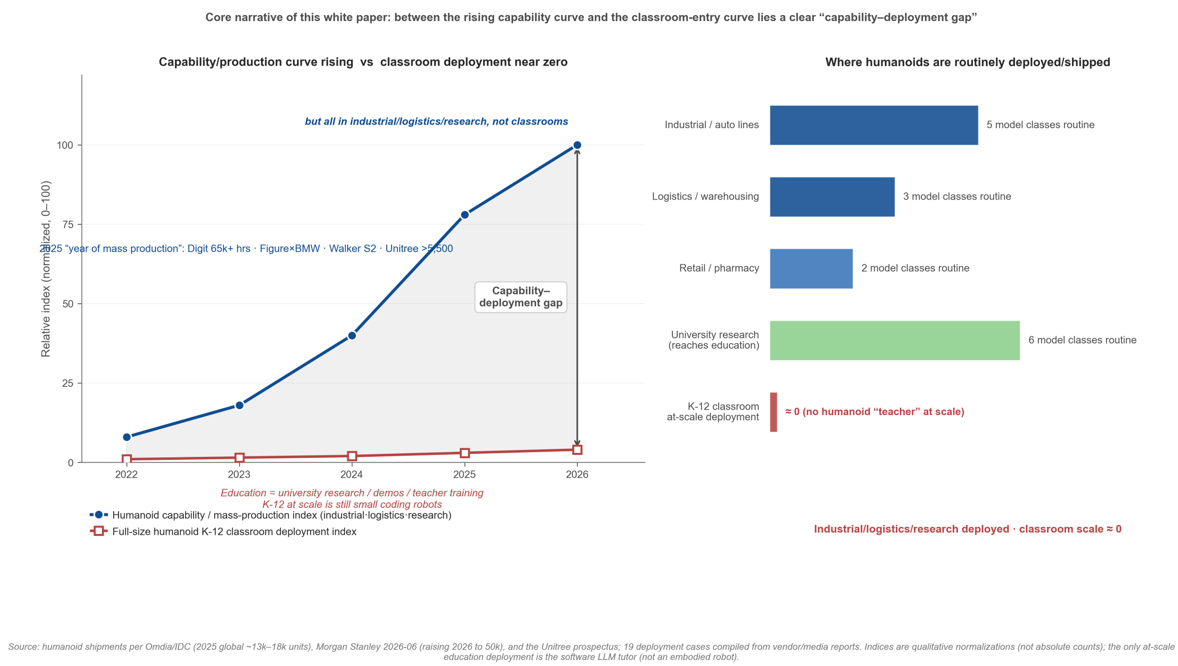


Figure 28.

This gap can be read at three levels:

- **The hardware trend is real**: humanoid DoF has risen to 20–56, with electric drive the mainstream; the dexterous hand's tactile density is led by Chinese vendors and its price beaten down to the 1/20 order of magnitude of overseas; the quadruped's motion control is engineered and its price has entered the thousands-of-dollars range. Cost continues to fall, pulled down by the Chinese supply chain. - **Mass production is almost entirely in non-education settings**: the humanoids genuinely in routine/at-scale commercial deployment (Digit's 65,000+ logistics hours, Figure 02's 30,000+ BMW vehicles, Apollo at Mercedes/GXO, Atlas with committed customers, Galbot at pharmacies, Walker S2's mass-production start) are **all in industry/logistics/retail**; those that are chiefly demonstration + claim are Optimus, Figure 03's chores, and 1X NEO's "autonomy" (in fact teleoperation). - **Educational accessibility is strictly tiered**: the humanoids that can enter education (Unitree G1, Fourier GR, Booster T1, EngineAI, Robot Era, Songyan) are **chiefly university/research teaching-and-research platforms, targets for**

teacher training, and competition platforms — Unitree's prospectus discloses that about three-quarters (73.6%–74%) of its humanoid revenue comes from research-and-education customers and only about 9% enters actual industrial applications ([107], [142]); Western humanoids (Figure/Tesla/1X/Apptronik) **barely enter education and serve only industrial production lines** ([143], [8]); **large-scale K-12 use is still small programming/social bodies of the UBTech Alpha Mini type**; and the embodied platform genuinely entering the education classroom "in the present tense" is the cheaper, more motion-control-mature **quadruped robot dog** (confined to vocational/university settings).

From this the chapter sets up its **red-line judgment**: the entry of full-size, industrial-grade humanoids into primary and secondary classrooms remains, to date, at the concept and demonstration stage; **"humanoid-robot teachers entering classrooms at scale" is an overclaim this blue book strictly avoids**. The body is maturing and cost is falling, but this first translates into landing in "industrial production lines + university research platforms + vocational practical training," not into "humanoid teachers entering basic-education classrooms."

The chapter's scope boundary bears restating: the above is about **body hardware and maturity** only. The market size and shipment forecasts of humanoids and quadrupeds belong to Chapter 12, vendor financials and competitive rankings to Chapter 13, the value-chain restructuring of upstream "bottleneck" components to Chapter 11, and specific educational-deployment cases to Chapter 9; they are mutual background to this chapter's hardware curves but are not developed here, to maintain cross-chapter de-duplication.

6.7 Chapter Summary

This chapter split embodied bodies into the three classes of humanoid, dexterous hand, and quadruped for examination one by one, landing on a cautious judgment: hardware is maturing and cost is falling, but the "year of mass production" production is almost entirely in industry/logistics/research, with humanoid landing in the education classroom near zero; the platform genuinely entering education (vocational/university) in the present tense is the cheaper, more mature quadruped robot dog, not the humanoid.

Following the thread that "once the body plugs in an LLM brain it will collect large volumes of multimodal behavioral streams," Chapter 7 turns to **multimodal interaction and learning analytics**: it analyzes the technology stack fusing speech, vision, affective computing, and LLM reasoning, discusses the value and limits of "embodied co-presence" as a robot's distinctive advantage, and warns, on emotion/attention monitoring, of the EU AI Act's emotion-recognition red line. The upstream-bottleneck specifics touched on in this chapter belong to Chapter 11, and red-line clauses such as emotion recognition to Chapter 14; this chapter offers only a pointer.

Chapter 7 Multimodal Interaction and Learning Analytics: Behavior Sensing, Controlled Research, and the Emotion-Recognition Red Line

> ***Overall maturity judgment:*** *the systems treated in this chapter are* ***almost all controlled research;*** *routine classroom deployment is scarce. Emotion and attention monitoring run straight into the EU AI Act Article 5(1)(f) emotion-recognition red line — this chapter issues a* ***global design no-go warning;*** *the clause-level details belong to Chapter 14.*

7.1 Introduction: The Perceive–Understand–Intervene Loop, and a Cautionary Thread Throughout

Chapter 6 left a thread at the hardware level: once a social body hosts an LLM brain and is placed in a real classroom, it **simultaneously collects a large multimodal behavior stream** — speech, expression, posture, and gaze. The question here is how to organize those streams into a "perceive–understand–intervene" loop: perceive the multimodal signals, understand the learner's engagement and emotional state, and make an adaptive intervention.

"Perceive–understand–intervene" is a **loop:** the perception side captures multimodal signals, the understanding side maps them into estimates of learner state (engagement, confusion, emotion), the intervention side adjusts teaching pace, prompts, or feedback accordingly, and the new signals after intervention re-enter perception — in theory, the faster and more accurate this loop, the finer the personalized teaching. Yet what this chapter argues is precisely that, in 2026, no link of this loop yet holds routinely in classrooms, and its most critical "understanding side" runs head-on into a regulatory red line.

Two cautionary lines run through the chapter. The first is **maturity:** the representative work cited below is, with almost no exception, controlled research, and routine classroom deployment is exceedingly scarce — fully consistent with the capability–deployment gap that Chapter 4 judged overall. The second is the **ethical red line:** collecting and analyzing emotion and attention runs directly into the EU AI Act's hard constraint banning emotion recognition in education institutions ([13]), and the chapter issues a warning on that point.

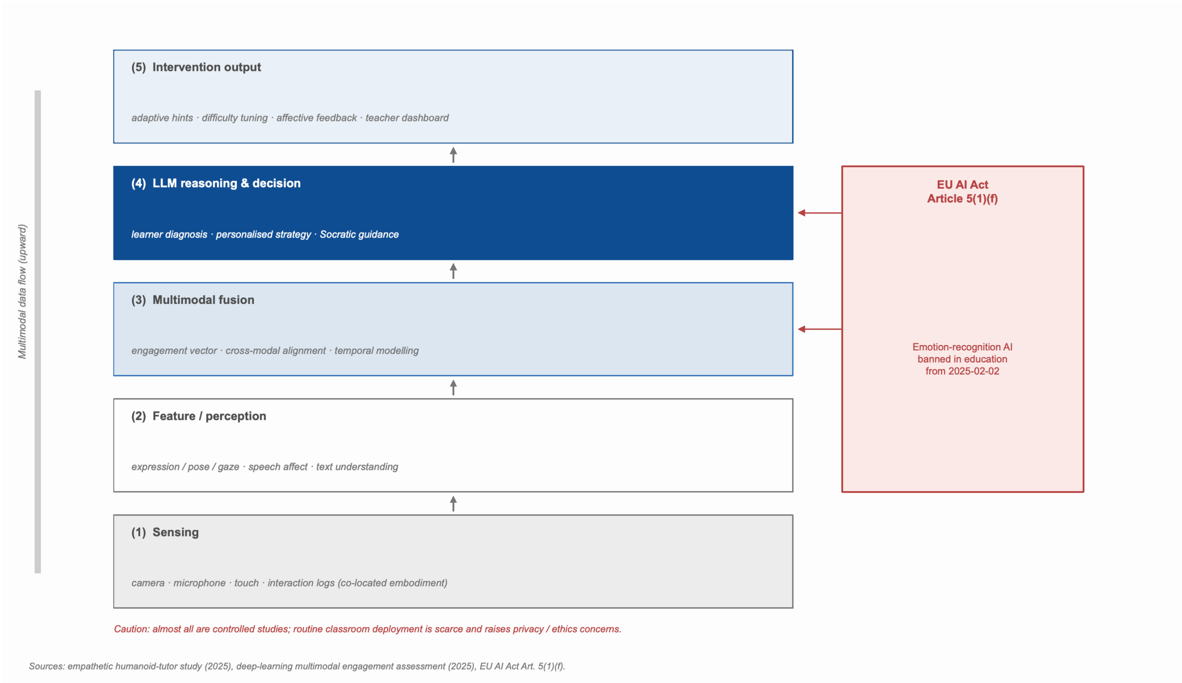


Figure 09.

7.2 The Technology Stack: Speech + Vision + Affective Computing + LLM Reasoning

The technology stack of multimodal learning analytics can be summarized as the capture of four classes of signal and a fusing layer of reasoning:

- **Speech:** recognizes semantic content and extracts paralinguistic cues from intonation, pauses, and speaking rate — the former answers "what the student said," the latter "in what state the student is saying it," both indispensable for judging comprehension and emotional state. - **Vision:** captures expression, posture, and gaze direction — the core channel for judging engagement, but also the channel of greatest privacy and ethical sensitivity, because it makes close-range, continuous recordings of children's faces and gaze. - **Affective computing:** maps the above signals into a state estimate of emotion and engagement — the most contested link in the stack, being at once the technical precondition for adaptive tutoring and the direct object of the §7.4 regulatory red line. - **LLM reasoning fusion:** uses the large model as the upper "understand–decide" hub, integrating the multimodal state and generating an adaptive response, converging dispersed perception signals into a single coherent teaching intervention.

Representative studies sketch the shape of this stack (each tagged for maturity; **almost all controlled research**):

System / work	Year	Description	Maturity
Empathetic humanoid (Llama 3.2 multimodal + emotion/memory/gesture, "Engagement Vector")	2025	A multimodal LLM drives empathetic interaction and builds an engagement vector	lab study (peer-reviewed, Frontiers; [144])

Effect of multimodal robots on engagement (Kebbi vs. Minibo comparison)	2024	Compares the effect of different social bodies' multimodal interaction on engagement	piloted / lab study ([145])
Multimodal deep learning to assess student engagement	2025	Uses deep learning to fuse multimodal signals and assess engagement	lab study (PLOS One; [146])
MathBuddy (face-api.js affective math tutoring)	2025	Affective math tutoring based on facial-expression recognition	lab study ([147])
Generative-AI social robots in education (Delphi expert panel)	2025	A Delphi method gathers expert judgment on GenAI social robots in education	research / policy review ([148])

The common ground of this table is that the technology stack is maturing fast — multimodal fusion plus LLM reasoning has become standard — yet every entry stays at the level of controlled research: they show that "this can be done," but not yet that "it already runs routinely in classrooms." Nearly half the work in the table uses **facial-expression / emotion recognition** as its core signal channel — exactly the object of the §7.4 red line — so the evolutionary direction of the stack and the regulatory no-go zone meet head-on.

Examining these representative works one by one reveals the realistic grade of the stack and its boundaries more clearly. The empathetic humanoid ("Engagement Vector") is built on the Llama 3.2 multimodal model, integrating emotion, memory, and gesture into a continuous "engagement vector" in an attempt to let the robot model learner state in real time and adjust its interaction accordingly (peer-reviewed, *Frontiers*; [144]); MathBuddy uses browser-side face-api.js directly for facial-expression recognition, feeding the recognized emotional state into a math-tutoring loop ([147]) — these two works demonstrate exactly how "affective computing" moves from concept toward a runnable prototype, but for that very reason they fall most directly within range of the EU emotion-recognition red line. Other work uses deep learning to fuse multimodal signals to assess student engagement (*PLOS One*; [146]), and a pilot study compares the effect of different social bodies' (Kebbi vs. Minibo) multimodal interaction on engagement ([145]). It must be stressed that each of these is a controlled lab or small-scale pilot, with generally limited sample size and duration, and that what they measure is mostly "engagement" rather than "learning gain" — a distinction that is the key caution of §7.3 and Chapter 10.

More broadly, a Delphi study gathering expert judgment ([148]) and a peer-reviewed third-party census covering 28 countries, 43 robot models, and 206 real deployment cases ([149]) together give a sober conclusion: **most social robots remain at the testing or pilot stage,** and routine deployment is the minority. This external anchor is fully consistent with Chapter 4's overall judgment that "the capability–deployment gap is not yet closed," and reminds us that the maturity of the technology stack far exceeds its degree of classroom uptake.

7.3 Embodied Co-Presence: A Distinctive Advantage of the Robot, and Its Limits

Relative to a purely software agent, the social robot has one distinctive advantage in multimodal learning analytics — **embodied co-presence.** Physical presence lets the robot simultaneously collect the learner's multimodal behavior stream: it can "see" the student's posture and gaze and "hear" the rise and fall of intonation, and it can respond with body posture, eye contact, and gesture. This is an interaction channel that a screen-based agent struggles to replicate, and the truly irreplaceable value of the "body" side in the brain–body separation of Chapter 5.

The other face of the advantage is a limit, and it must be presented honestly. **The stronger the collection capability, the greater the privacy and ethical risk** — embodied co-presence means the close-range, continuous collection of children's sensitive signals such as facial data, voiceprints, and gaze (touching directly on the red line of §7.4). The tension is inherent: what gives the embodied robot its distinctive value is exactly its ability to capture the learner's multimodal behavior stream up close and continuously, and that same ability makes it the most concentrated carrier of risk for collecting children's sensitive data. Faces, voiceprints, and gaze trajectories are all strongly identifying biometric signals; collected continuously in a classroom, their compliance and ethical burden far exceeds that of ordinary homework data — collection capability and compliance risk are two sides of one coin, and one cannot take the benefit while ignoring the harm. **A rise in engagement is not a learning gain** — multimodal signals most often measure engagement, but rising engagement need not translate into a cognitive learning gain, a distinction Chapter 10 treats systematically and in evidence-based terms (engagement ≠ learning gains is one of that chapter's key cautions). Embodied co-presence lets the robot "measure more," but "able to measure," "able to measure accurately," and "beneficial to learning" are three different things.

Placing the "perceive–understand–intervene" loop within the real triad of teacher, student, and robot brings the limits further into view. The ultimate purpose of learning analytics is not to generate a handsome engagement dashboard but to return trustworthy learning information to the **teacher** in support of their judgment and decisions; the reasonable positioning of a robot in the classroom is mostly that of a teacher's assistant and a mediator of learning, not an independent arbiter that replaces the teacher. This means that the effectiveness of a multimodal learning-analytics system depends not only on "measuring accurately" but, more so, on "whether the measured information can be reliably interpreted and adopted by the teacher" — a system that pushes a noisy emotion estimate to a teacher as a hard indicator may be worse than no system at all. One more echo of Chapter 6 must be added: the value of embodied co-presence rests on the body being able to reliably "be present, perceive, and respond"; and Chapter 6 has shown that the embodied bodies able to enter education remain mostly university/research platforms and quadruped training platforms, while multimodal robots that are routinely co-present in primary and secondary classrooms do not exist. Thus "embodied co-presence" in 2026 is more a **validated research advantage** than a classroom capability already rolled out at scale — it "measures more," but "able to measure," "able to measure accurately," and "the measured information benefits teaching" are three things that must be proven separately.

7.4 A Privacy and Ethics Red-Line Warning: Echoing the EU Emotion-Recognition Ban

The most sensitive link in multimodal learning analytics is the **recognition of emotion and attention** — and this is exactly where it strikes the clearest global regulatory red line at present.

EU AI Act Article 5(1)(f) **prohibits the use of AI to infer the emotions of natural persons in education institutions** (emotion recognition), one of the "prohibited AI practices" under Article 5, applicable since February 2, 2025 ([13]). This means that classroom robots and camera systems carrying "attention monitoring" or "emotion recognition" functions are **already directly restricted** in EU educational settings — this is not a grey area under the high-risk category that "may be used after a conformity assessment," but an expressly **prohibited** no-go zone. One legislative state of affairs, easily confused, must be clarified: although simplification proposals such as the EU Digital Omnibus had not yet been finally adopted as of June 2026 (a provisional political agreement plus parliamentary endorsement, with the formal text pending publication), **the Article 5(1)(f) emotion-recognition ban is already in force and is unaffected by the simplification process** — drafting and product judgment must treat this ban as an established constraint, not soft-pedal it on the grounds that "the legislation is still in flux."

Combined with the observation in §7.2 — that nearly half the representative multimodal learning-analytics systems use facial emotion recognition at their core — this red line is clearly not a marginal constraint but one that points straight at a mainstream path of the current technology stack. In other words, the design directions most favored by researchers today — "engagement vectors," "attention heatmaps," "emotion-adaptive tutoring" — are exactly the first to be drawn out of bounds by law in EU educational settings. The arrow of technical evolution and the arrow of regulatory prohibition collide head-on, which is precisely why this chapter places it at the core of the cautious sections.

On this basis the chapter issues a **global design no-go warning:** given the directionality of this ban and the trend of regulatory convergence, "classroom emotion recognition" should be treated by product designers as a **global design no-go to be actively avoided,** not a compliance item to be avoided only in the EU. Regulatory convergence is no abstraction — beyond the EU ban, China's Measures for Labeling AI-Generated and Synthetic Content and its anthropomorphism-related rules, and the US COPPA's inclusion of biometric information within the scope of children's personally identifiable information (PII) protection, are all tightening, from different directions, the constraints on collecting children's sensitive signals (clause-level details consolidated in Chapter 14). Against a backdrop of multiple jurisdictions tightening in the same direction, making "emotion recognition" a core selling point is, on compliance, an ever-narrowing road.

Two boundaries warrant care. This chapter issues only a warning and a pointer; the **specific clauses, scope of application, and relationship to other jurisdictions are consolidated in Chapter 14** and not repeated here. And the warning targets the class of function that "recognizes emotion and intervenes on that basis," not multimodal interaction itself — a robot may still, without inferring emotion, use speech and vision for harmless interactive assistance. Abundant **compliant design alternatives** exist: using student-initiated questions and answers, rather than passive emotion monitoring, as the interaction trigger; replacing automatic emotion intervention with explicit teacher-in-the-loop feedback; and confining perception strictly to the minimum necessary to complete the task, with data processed locally and not transmitted out. These paths sidestep the red line while retaining the pedagogical value of multimodal interaction.

Finally, the ethical red line must be set beside the outcomes red line of Chapter 10: even if emotion/engagement recognition holds up technically and on compliance, "a rise in engagement" need not equal "a learning gain" — the Netherlands' L2TOR project, a multi-site study of 194 children around 5 years old, found that a NAO robot plus a tablet teaching second-language vocabulary held **no significant advantage** over the tablet alone ([150]). This near-zero-gain finding is a reminder: the more multimodal data is collected, the greater the need to guard against misreading "measured engagement" as "what was learned." The convergence of the chapter's three cautionary lines is thus visible — technically, "almost all controlled research, with routine deployment scarce"; ethically, "emotion recognition already drawn inside the red line"; and on the evidence, "engagement ≠ learning gains" — and together they fix the real-world positioning of multimodal learning analytics in 2026 as **cautious exploration, not mature deployment.**

7.5 Chapter Summary

This chapter has drawn the "perceive–understand–intervene" technology stack of multimodal interaction and learning analytics: the fusion of speech, vision, affective computing, and LLM reasoning has become standard, and embodied co-presence is the robot's distinctive collection advantage. But two cautionary lines run through the whole chapter — first, the representative work is **almost all controlled research,** with routine classroom deployment scarce; second, emotion and attention recognition have already struck the EU AI Act's red line, and the chapter treats "classroom emotion recognition" as a global design no-go warning. We have also distinguished a "rise in engagement" from a "learning gain," laying an interface for the evidence-based discussion in Chapter 10.

To close the chapter's real-world positioning in one sentence: the **technical maturity** of multimodal learning analytics is markedly higher than its classroom-uptake maturity, and higher still than its compliance feasibility. The stack can already fuse speech, expression, posture, and gaze into a coherent learner-state estimate, and research prototypes proliferate; but turning that capability into a routine, compliant, and genuinely learning-effective product in primary and secondary classrooms still faces three uncrossed thresholds — scarce evidence of routine deployment, a clear legal no-go zone for emotion recognition, and the evidence-based caution that "engagement ≠ learning gains." The EU emotion-recognition red line this chapter flags is settled at the clause level in Chapter 14; the "engagement ≠ learning gains" interface this chapter lays down is honored systematically, with effect sizes, in Chapter 10.

Part Two (technology) has now characterized the four facets of evolution, agentification, the body, and multimodality, which together support a single judgment: **frontier capabilities are already rich, but**

deployment remains thin; the brain has landed, but the embodied body and the multimodal loop have not yet landed in the classroom. Taking up this chapter's threads — "engagement ≠ learning gains" and "controlled evidence for embodied LLM robots is scarce" — Part Three turns to **applications:** Chapter 8 lays out the full spectrum of application scenarios and market calibers, Chapter 9 maps the vendor landscape distinguishing "deployed" from "demonstrated/claimed," and Chapter 10 presents, scenario by scenario and by strength of evidence, the systematic evidence on learning outcomes — recording null and adverse effects with equal prominence. The EU emotion-recognition red line that this chapter flags is settled, at the clause level, in Chapter 14.

Part III Deployment and Evidence: Who Actually Reached the Classroom

The full application spectrum · real cases and pilots · an evidence-based look at learning outcomes

This part answers the decisive question — who actually reached the classroom, and with what effect: from the full application spectrum, through evidence-graded cases and pilots, to scenario-by-scenario evidence on learning outcomes. It keeps a strict line between deployed and merely demonstrated, strong evidence and weak, with null and adverse effects equally recorded.

Chapter 8 The Full Spectrum of Application Scenarios

Application Spectrum — Maturity and Market Sizing across Five Sub-tracks

---

8.1 An Overview of the Spectrum: Five Tracks from the Most Mature to the Most Promising

Part Two (Chapters 4–7) asked what educational robots can do in 2026 and how far the underlying technology has matured. This chapter asks a question that sits closer to the working life of educators and industry decision-makers: **where, exactly, do these capabilities land in real educational settings, and what is the maturity and the credible market sizing of each setting?** This is the decisive step from the technology-supply side to the application-demand side, and it opens Part Three (Application).

Continuing the **capability–deployment gap** that runs through the entire white paper, this chapter holds to one discipline as it describes each track: **maturity and market sizing are flagged separately, and where credible cross-source figures are missing we leave a deliberate blank (PENDING) rather than force a number.** This is of a piece with the honest retrospective in Chapter 12, which revisits the field's historical forecasts (the 2016 and 2019 editions projected US$11.1 billion for 2021 and US$84.1 billion for 2023, whereas the 2026 measured market — even on the most aggressive scope — is only about US$5.8 billion). The greatest epistemic risk in the educational-robot market has never been underestimation; it is the overestimation that comes from mistaking a broad-scope consumer market for the narrow-scope educational-robot hardware market.

We group the application scenarios of educational robots into **five sub-tracks**, positioning each along three dimensions — maturity, strength of evidence, and availability of credible market sizing ():

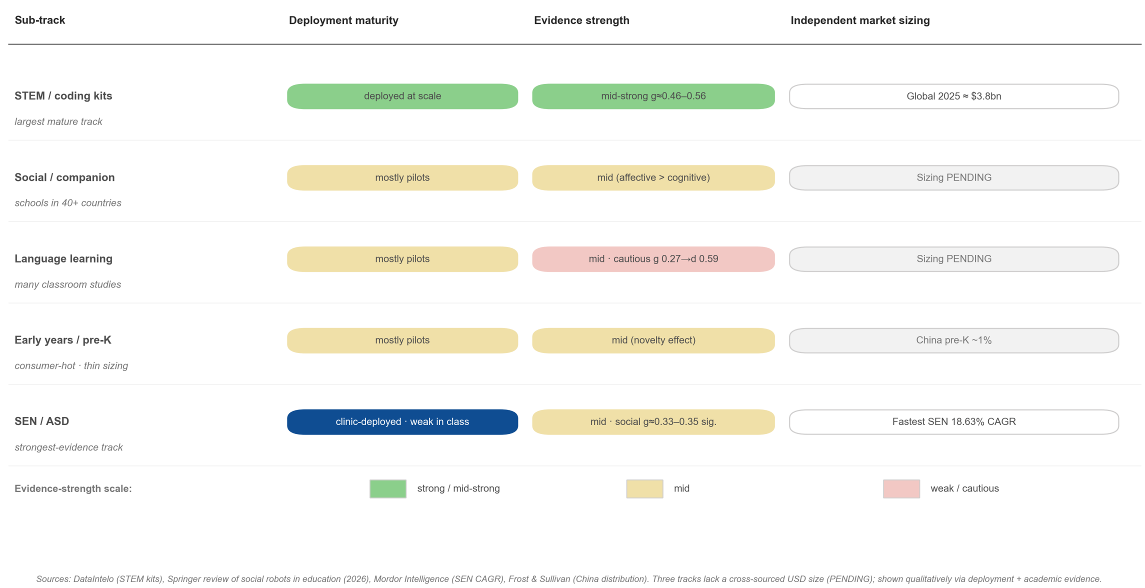


Figure 10.

Track	**Maturity**	**Strength of evidence (see Chapter 10)**	**Independent dollar market sizing**
① STEM / coding kits	Largest, most mature, already at commercial scale	Medium positive effect; relatively the most robust	Cross-source figures available (≈ US$3.8 billion / 2025)
② Social / companion (teaching assistant · peer tutor)	Research and pilots; routine deployment scarce	Affective gains > cognitive gains	PENDING (no credible cross-source figure)
③ Language learning	Research and pilots	Medium, but the conservative conclusion must be set alongside it	PENDING
④ Early childhood / preschool	Hot consumer market, small share within the statistical scope	Affective > cognitive; high novelty-effect risk	PENDING (only ≈ 1% of the scope)
⑤ Special education / inclusion (ASD)	Fastest-growing; strongest yet most contested evidence	Strong, but cross-setting extrapolation is in doubt	Single-firm sizing (special-education CAGR 18.63%)

Two sizing disciplines are made explicit at the outset and held throughout the chapter.

Track boundaries and statistical scope are tightly coupled. STEM and coding kits have a relatively clear dollar size because they correspond to countable hardware shipments (kits, controllers, sensor packs). The social/companion, language-learning, and early-childhood tracks, by contrast, carry much of their value in software subscriptions, content, and services. As §8.3 shows, hardware accounts for only 50%–60% of the total cost of ownership (TCO) of a social robot [151], with the remaining 40%–50% falling to operations, training, and software subscriptions — value the conventional "robot-shipment" lens cannot capture. That is the structural reason an independent dollar size for these three tracks remains PENDING; it is not an oversight in the research.

The robotics-service scope and the robot-hardware scope must never be conflated. The track-level figure given here — robotics-education services at roughly US$2.28 billion in 2026, rising to US$11.6 billion by 2034 [152] — is a **service-scope** figure, belonging to a different statistical system from the global educational-robot **hardware scope** of Chapter 12 (roughly US$2.0–2.8 billion in 2026). The two are presented side by side, each labeled, and must never be summed or interchanged. This chapter draws only the track-level portrait; the **global market TAM is reconciled in Chapter 12** (de-duplicated across chapters).

8.2 STEM / Coding Kits: The Largest and Most Mature Track

STEM and coding kits are the **only sub-track that can be called "deployed at commercial scale,"** and they correspond directly to "Educational Robotics" in the binary framework of Chapter 3 — kit-based learning in which the learner assembles and programs the robot. Their technological base is not the frontier foundation models of 2026 but the line of **pre-foundation-model programmable kits**: LEGO Mindstorms (EV3), VEX, mBot, Thymio, and the like. Brick-based construction paired with block-based programming cultivates computational thinking through deterministic, rule-based behavior [51]. LEGO discontinued the Mindstorms line — the emblematic product of this paradigm — in 2022, a move that signals not a shrinking track but a generational shift of block-based coding kits toward lighter, more software-centric forms.

Market sizing (cross-source, S1). The global market for STEM robotics kits stood at roughly **US$3.8 billion in 2025**, with a compound annual growth rate (CAGR) in the **5.5%–7.8%** range, cross-sourced by Dataintelo [153] and Industry Research [154]. Two scope details deserve attention. The two sources disagree on the CAGR (5.5% versus 7.8%), and this blue book presents both rather than selecting a single value. And the 5.5%–7.8% growth rate runs **markedly lower** than the 15%–29% range projected for the global educational-robot hardware market in Chapter 12, because the STEM-kit segment is a relatively mature, installed-base market whose growth momentum lags the emerging hardware tracks driven by the humanoid and embodied-intelligence narratives. That contrast — moderate growth for the mature track, aggressive growth for the emerging ones — is itself an important signal about the structure of the educational-robot industry.

Why it is the most mature. On the evidence side (see Chapter 10, §10.2), STEM, coding, and computational thinking constitute the **most robustly evidenced** of all application scenarios: several independent meta-analyses converge on a medium positive effect (g / SMD ≈ 0.46–0.56), and at least one computational-thinking meta-analysis found no marked publication bias on Egger's test [155]. This means that, among the five tracks, the alignment between market size and educational value rests on the **strongest empirical support** here — the one track confirmed in both directions, by industry and by evidence. The point matters for industry decisions: in a field saturated with "promising but early-stage" narratives, STEM and coding are among the few tracks with both a countable market and credible evidence of effect, so the evidential risk in their investment and procurement decisions is comparatively the lowest.

Representative vendors and the competitive landscape. These include LEGO Education (the long-standing leader), VEX, Sphero, Makeblock, Wonder Workshop, ROBOTIS, and UBTech. This is the **most fully contested track, with the best-developed export channels,** in the educational-robot industry. Chapter 9, §9.6 details the global channel build-out of Dobot (6,500+ institutions across 80+ countries) and Makeblock (140+ countries, more than 20 million users) — the strongest evidence of this mature track's "already at commercial scale" character. It should be stressed that, although the track's technological base remains the pre-foundation-model programmable kit, the foundation models and embodied capabilities of 2026 are entering through an "upgrade of the programming target": students' programming aims are expanding from block-based robots toward conversational agents and embodied platforms. This evolution is taken up further in the value-chain restructuring of Chapter 11.

8.3 Social / Companion and Language Learning: Rich in Qualitative Detail, Independent Sizing To Be Completed

Social/companion robots and language learning are two tracks that are **rich in academic research and clear in their deployed hardware, yet have no credible cross-source figure for an independent dollar market size.** This section follows the discipline established in §8.1: it presents each track qualitatively through deployed hardware and academic evidence, and does not force a number.

Social / Companion (Teaching Assistant · Peer Tutor)

This track corresponds to the core of the "Educational Service Robots" category in the binary framework of Chapter 3 — service robots with teaching-and-learning intelligence. Academic reviews indicate that schools in **more than 40 countries** already use humanoid and social robots as teaching assistants, peer tutors, or special-education support [156]. The principal classroom platforms include NAO, Pepper (the two most common, though Pepper was discontinued in 2021; see Chapter 6), QTrobot, UBTech's Alpha Mini, and Furhat, at unit prices ranging from US$2,000 to US$30,000.

The root reason an independent dollar size remains PENDING for this track lies in its cost structure: **hardware accounts for only 50%–60% of the total cost of ownership (TCO)** [151] (a single-media figure, usable for a qualitative read of cost structure), with the remaining 40%–50% sitting in operations, teacher training, course content, and software subscriptions. Chapter 11 distills this cost structure further into a value-chain judgment — that operations and curriculum certification have been elevated into an independent value layer. For this chapter, it explains why the conventional "shipments × unit price" lens cannot credibly portray the track: much of the value flows to a service side that hardware counts cannot capture.

Language Learning

Language learning is one of the applications of social robots that has drawn the most academic attention. A repeatedly observed phenomenon is that **a student's anxiety when speaking a foreign language to a robot is markedly lower than when speaking to a person, and engagement rises accordingly** [151] — and this positioning as a "low-social-pressure practice partner" is the distinctive value of robots in language learning. Robot-Assisted Language Learning (RALL) already rests on a documented literature base, including Randall's (2019) review of 79 studies [157].

Here, however, the blue book must strictly enforce **evidence firewall E2**: the language-learning track may **never be said to be "proven effective."** The authoritative review by van den Berghe et al. (2019) reaches a conservative conclusion — robots cannot yet be confirmed as effective language tutors [158] — and the existing evidence generally shows that **affective and motivational gains outweigh cognitive gains**, with the overwhelming majority of studies being single-session, small-sample work in which the novelty effect cannot be ruled out. The full effect-size grading and the conservative argument appear in Chapter 10, §10.3; at the level of the track portrait, this chapter offers only a qualitative position: **deployed hardware is clear, affective value is clear, cognitive effects are treated cautiously, and an independent dollar size remains PENDING.**

8.4 Early Childhood / Preschool: High Consumer Heat, Small Share within the Statistical Scope

Early childhood and preschool is a deceptively positioned track: **consumer demand runs very hot, yet its share within the strict statistical scope of "educational robots" is very small.** This contrast is the most concentrated expression of the risk emphasized in §8.1 — that a broad-scope consumer market is not the narrow-scope educational-robot hardware market.

Take the Chinese market. On Frost & Sullivan's 2023 scope, when China's educational robots are distributed by educational stage, the **early-childhood stage accounts for only about 1%** [159]. Representative products such as the early-childhood robots of Toycloud (Alpha Egg) and iFLYTEK are highly visible in the public consumer eye, yet occupy a very small share within the strict "educational-robot" industry statistics — much early-childhood smart hardware is classified as consumer electronics or smart toys rather than educational-robot hardware.

This phenomenon carries two implications. First, on market sizing, the early-childhood track's **independent dollar size remains PENDING**: visible consumer heat cannot be credibly converted into a narrow-scope educational-robot market size. Second, on the evidence side (see Chapter 10, §10.4), the early-childhood track is precisely the scenario with the **highest novelty-effect risk** — young children are especially sensitive to novel stimuli, and a short-term rise in engagement is easily misread as long-term learning gain. The white paper holds to an "affective gains > cognitive gains" position for this track and defers the evidential detail to Chapter 10.

8.5 Special Education / Inclusion: Fastest-Growing, Strongest Evidence, Most Contested

Special education and inclusion — exemplified by intervention for autism spectrum disorder (ASD) — occupies a singular place among the five tracks: it is **simultaneously the fastest-growing, the most strongly evidenced, and the most contested.**

Market sizing (single firm, S3). Mordor Intelligence describes the special-education segment as the fastest-growing, with a CAGR of **18.63%** [160]. This must be flagged explicitly as a **single-firm figure** from Mordor, with no second credible source for cross-verification; cite it with caution.

Why the evidence is the strongest. From an evidentiary standpoint (see Chapter 10, §10.5), ASD intervention is the **only track supported by a substantial body of RCT meta-analyses** in the educational-robot field. The most robust is the PLOS One (2022) meta-analysis: a review of 40 studies and a meta-analysis of 12 RCTs across 346 participants, with an overall Hedges' $g = 0.33$ (95% CI [0.08, 0.57]) and a **significant *g* = 0.35 on the social-functioning dimension** [161]. In terms of literature concentration, NAO appears in roughly 39% of ASD-related robot studies, making it the de facto standard platform for the track [53] [162].

Why it is the most contested. It is precisely this most strongly evidenced track that is the most easily overstated. The white paper flags **evidence firewall E3** here in advance: ASD intervention must **never generalize "clinic-effective" into "universally effective in classrooms or homes."** The setting-moderation analysis within the same PLOS One meta-analysis shows that the clinic setting yields a significant $g = 0.57$, but the **school setting yields *g* = −0.16 and the home setting *g* = 0.16, both non-significant**; the emotion dimension ($g = 0.63$, but with an extremely wide CI of [−1.43, 2.69]) and the motor dimension ($g = -0.10$) show null or non-significant effects; and only 17.5% of the studies were rated "strong" in quality [161].

This tension — the strongest evidence yet the most constrained extrapolation — is the defining feature of the special-education track: it has the most RCT support in the academic literature and the fastest projected growth in the market (18.63%), yet its genuine significant effect is largely confined to the clinic, a controlled environment, rather than the school classroom where real educational deployment takes place. Treating clinic evidence as a classroom endorsement is the most dangerous epistemic leap in this track. The detailed strength-grading, the setting-moderation analysis, and the equal recording of null and adverse effects are the core task of Chapter 10, §10.5; at the level of the track portrait, this chapter offers only the position "fastest-growing, strongest evidence, but extreme caution required in cross-setting extrapolation."

8.6 Chapter Summary

Using the three-dimensional frame of maturity, strength of evidence, and availability of market sizing, this chapter has drawn a full-spectrum portrait of the five educational-robot sub-tracks (Fig. 10), which can be distilled into three judgments.

First, **maturity is distributed as a gradient.** STEM and coding kits are the only track already at commercial scale (≈ US$3.8 billion / 2025 [153]); social/companion and language learning are dominated by research and pilots; early childhood is hot in consumer terms but small within the statistical scope (the early-childhood stage at only about 1% [159]); and special education is the fastest-growing (18.63% [160]) but the most contested.

Second, **the sizing discipline cannot be relaxed.** The social/companion, language-learning, and early-childhood tracks have **no credible cross-source figure for an independent dollar size (PENDING)**; this blue book presents them qualitatively through deployed hardware and academic evidence, and fabricates no numbers. The robotics-service scope (≈ US$2.28 billion / 2026 [152]) and the hardware scope of Chapter 12 belong to two separate systems and must not be conflated.

Third, **evidence and market must be stated on separate tracks.** Market heat (above all in early childhood and special education) is not evidential strength; this chapter offers only a qualitative positioning of evidential strength at the track level, and hands off all effect sizes, null/adverse-effect records, and cross-setting extrapolation risks to Chapter 10 for strict scenario-by-scenario grading.

A look ahead. Chapter 8 has answered which tracks exist and what their respective maturity and sizing are; **Chapter 9** enters the real-world deployment map of these tracks, distinguishing "deployed" from "demonstrated/claimed" case by case — from the primary- and secondary-school-scale deployment of UBTech's Alpha Mini, to the university-research-platform positioning of Unitree's G1 EDU, to the reality that Western humanoids are "industrial only, with no education" — mapping vendor deployment onto an evidence-tier grid. **Chapter 10** then takes up the "strength of evidence" flags from each track in this chapter and develops the full learning-outcomes evidence, grading effect sizes and strength by scenario and recording null and adverse effects with equal prominence.

Chapter 9 Cases and Pilots: An Evidence-Graded Map of Real-World Deployment

> *Cases and Pilots — A Nineteen-Case Evidence-Graded Map of Who Entered Education, and to Which Layer*

9.1 Method: A Dual Yardstick of Educational-Entry Status and Evidence Tier

Chapter 8 portrayed the maturity and market sizing of the five sub-tracks; this chapter brings the lens in close, to **the real deployment status of specific vendors, products, and policies.** This is where the blue book's central spine — the capability–deployment gap — most needs empirical testing. Industry news is saturated with grand narratives of "robots entering classrooms" and "a hundred cities, ten thousand schools," and this chapter's task is to sort those narratives, one by one, with a cool dual yardstick. The chapter assembles **19 real-world deployment cases** (10 in China, 9 international), each carrying a scale figure and a maturity tag — the most systematic empirical stocktaking of the deployment side in this blue book.

For every case, this chapter flags two dimensions at once, consistent with the four-tier maturity-label system established in Chapter 4:

- **Educational-entry status:** whether, and how, the product has entered a real educational setting (primary/secondary classroom / university research platform / planning stage / no educational deployment); **- Evidence tier:** the strength of the evidence behind that status (`[deployed]` — open commercial supply, government documents, or independent on-site reporting / `[piloted]` — small-scale real settings / `[demonstrated]` — controlled evaluation or vendor video / `[vendor-claimed]` — company figures, not independently verified).

Three cross-chapter de-duplication disciplines frame the discussion. The chapter **addresses deployment status only, not financials or shipment rankings** — UBTech's revenue, Unitree's valuation, iFLYTEK's education revenue and other financial/competitive data belong to Chapter 13, and when a vendor is cited here it is drawn on only for the "educational deployment" side. Nor does the chapter **restate hardware specifications** (degrees of freedom, cost, harmonic drives and the like belong to Chapter 6) or recompute market size (which belongs to Chapter 12). And **wherever a Chinese vendor's financial figures appear, they are stated in RMB 亿元 (100 million yuan), never as billions —**

UBTech's 2025 AI-education revenue was RMB 413 million (about US$57 million), decidedly not "4.13 billion" [163]. The sole focus of this chapter is **who has genuinely entered education, to what layer, and how hard the evidence is.**

The central cautionary conclusion that runs through the chapter can be stated up front (humanoid-in-education, cooled — correction point 2): **what truly enters classrooms at scale is always one of two things — (i) screen-based software AI tutoring agents (Khanmigo / Amira / Gemini), and (ii) the pre-foundation-model generation of small coding/social robots (UBTech's Alpha Mini class, NAO / Pepper); the humanoids that can enter education are, for roughly three-quarters of the relevant revenue, university research and teaching platforms; Western humanoids barely enter education and serve industry only. Any "humanoid robot teachers at scale in classrooms" claim is strictly prohibited.** The full vendor-deployment picture appears in .

Figure 11.

Why insist so strictly on the distinction between "deployed" and "claimed"? Because the narrative ecology of the educational-robot industry has a structural feature: **vendors have a strong incentive to package "plans," "demonstrations," and "pilots" as "already deployed,"** and media coverage tends to amplify grand figures ("nearly 2,000 schools," "300 humanoids," "10 million students"). Unless each such claim is anchored to an evidence tier, the blue book would degenerate into a paraphrase of industry promotion. This chapter therefore asks three questions of every "deployment" claim: is there open commercial supply, an official government document, or independent on-site reporting (deployed)? Or a small-scale real setting (piloted)? Or merely a controlled demonstration or a company figure (demonstrated/claimed)? Only claims that survive these three questions are classed as "deployed."

A further policy boundary runs through the Chinese cases and is worth flagging at the outset — **"banned in examinations, advanced in teaching":** China's AI-in-education policy simultaneously pushes hard, hour-by-hour, on the teaching side (Beijing requires no fewer than 8 class hours per academic year; Shanghai makes the subject compulsory in grades 4 and 7) while banning generative AI from answering questions directly on the examination side (primary pupils must use it under teacher or parental guidance, and copying AI-generated content into answers is strictly prohibited [164]). When this chapter presents the class-hour mandate cases, it preserves this boundary and does not misread "teaching-side entry into the classroom" as "examination-side liberalization."

9.2 Chinese Cases (10): From Primary/Secondary Scale Deployment to University Research Platforms

The overall picture of the Chinese cases confirms the judgment of Chapter 8: **the most mature, most verifiable "deployed" cases cluster in three places — small coding robots entering primary and secondary schools, the hard class-hour mandates of AI general-education courses, and the large-scale expansion of embodied disciplines in universities; and where humanoids enter education, they enter the university research-platform layer, not the primary/secondary teaching layer.**

Case 1 | UBTech AI-Education into Schools — the primary/secondary scale-deployment mainstay [deployed]

UBTech (UBTech Robotics; HKEX: 09880) is the vendor in China's educational-robot field that comes **closest to meeting the definition of "deployed at scale."** Its AI-education business now serves teachers and students at **nearly 2,000 public primary and secondary schools** across China (three in-school formats: AI labs, AI general-education courses, and after-school extended-day courses), covering **25 provinces** domestically and **more than 20 countries** abroad; its 2025 educational-intelligent-robot business generated revenue of **RMB 413 million (up 13.7% year over year)** [165] [163] [166]. What genuinely reached primary and secondary classrooms at scale, however, is **Alpha Mini (Wukong) and similar small coding/social robots** — situated in the overlap of the STEM/coding and social tracks of Chapter 8, and the category with the most solid "deployed" evidence.

The scope flags must be honest: **"nearly 2,000 schools" is a vendor/annual-report figure** (a search-engine paraphrase, with no audited breakdown seen), and **"5,000 schools" is a target, not an accomplished fact** [165]; the RMB 413 million is an audited-annual-report figure and is **in 亿元 (100 million yuan), not billions.** This stands in sharp contrast to the reality of humanoid robots: what runs at scale in primary and secondary classrooms is the desktop-class small robot, not the full-scale humanoid.

Case 2 | UBTech Tien Kung Walker — university research-and-education humanoid [deployed (university research) + vendor-claimed]

In 2025, UBTech and the Beijing Humanoid Robot Innovation Center jointly released a full-scale research-and-education humanoid, the **Tien Kung Walker, priced at RMB 299,000** (comparable models generally exceed RMB 500,000); orders surpassed 100 units within four months, and **deliveries to the education-and-research field alone are projected to exceed 300 units in 2025**; the company has reached "joint humanoid-robot research and co-creation" intent agreements with Wuhan University, the Chinese University of Hong Kong, Zhejiang University, Shanghai Jiao Tong University, and others [167] [168].

The blue book grades this evidence separately: **the units already delivered into universities are "deployed (university research platform)," whereas the "300+ units" is a company projection and thus "vendor-claimed,"** and the university partnerships are "intent." The crux is that **the Tien Kung Walker is aimed at university research, not at primary and secondary classrooms.** Presenting the same vendor's "deployed small robots in primary/secondary schools" (Case 1) and "research humanoid into universities" (Case 2) side by side is precisely the point of this chapter's dual-yardstick method: **the same vendor holds both the most solid "deployed" case and an unverified "claim," and the two address entirely different education levels and must be classed separately.**

Case 3 | Unitree G1 (EDU) + RoboUniversity — the university-research mainstay [deployed (university research)]

The Unitree sample proves the other half of this chapter's core judgment: **the humanoids that genuinely enter education are almost all university and research teaching platforms.** In 2025 Unitree (Unitree Robotics) shipped **more than 5,500 humanoids (first in the world, a 32.4% share),** of which **roughly 74% were sold to universities for research and only about 9% entered actual industrial use;** its cumulative quadruped sales exceeded 30,000 units; labs at MIT, Stanford, CMU, and ETH Zurich use the G1 EDU; and Unitree, together with Stanford professors and OpenMind, co-runs the **RoboUniversity** curriculum-and-certification system [107] [169] [170].

The "roughly 74% sold to universities for research" (prospectus basis) and the source's "73.6% from research and education" (the January–September 2025 window) are presented side by side, and this blue book renders them uniformly as **"roughly three-quarters from research and education."** This figure is the strongest empirical support for the judgment that "the humanoid that can enter education = a university research platform" — the genuine buyers of humanoid robots in education are research labs, not primary- or secondary-school classrooms.

The essential difference between a "research platform" and a "teaching application" repays a closer look. In a university lab, the G1 EDU is **an object of study and a carrier for development**: doctoral students use it to run experiments in motion control, reinforcement learning, and human–robot interaction, and it functions as research infrastructure, akin to a precision instrument in the lab. In the envisioned "teaching application" of a primary or secondary classroom, by contrast, the humanoid is expected to become **a teaching agent facing students** — lecturing, answering questions, accompanying. The former is already a verifiable reality, with roughly three-quarters of revenue coming from research and education as the hard proof, whereas the latter, in 2026, remains at the level of demonstration and aspiration. RoboUniversity is likewise, in essence, a research/teaching-platform partnership aimed at universities, not a dispatch of humanoids to primary- and secondary-school classrooms as teachers. Seeing this distinction clearly is the key to resisting the "humanoid teachers in the classroom" narrative.

Case 4 | The Central Audio-Visual Education Hall's AI Curriculum and National Platform — a national resource base [deployed]

The Ministry of Education's Center for Educational Technology and Resource Development (the Central Audio-Visual Education Hall) has announced its **fourth batch of "Central-Hall Artificial Intelligence Courses" and "Central-Hall Virtual Experiments";** jointly with the MOE's Department of Basic Education it released the **"Yu Xiaomiao" agent** on the National Smart Education Platform for primary and secondary schools; and the national platform now aggregates **more than 1,000** flagship educational-AI courses across four categories [171] [172]. The very phrase "fourth batch" attests to years of batched rollout. **Maturity: deployed (national resource platform + course certification).** Note that this is a national-scale deployment at the "course and resource" layer, which is a different matter from "robot embodiments entering the classroom" and must not be conflated.

Case 5 | 509 National AI-Education Base Schools — a policy-grounded school network [deployed]

The Ministry of Education has established **509 artificial-intelligence-education base schools** nationwide to lead the way, and has selected **7 eastern provinces, 20 central and western prefecture-level cities, and 18 universities** to pilot the AI-empowered-education initiative [173] [174]. **Important calibration:** the "509 base schools" and the "7 provinces / 20 prefectures / 18 universities" come from a press briefing rather than the policy text proper, and since the 20 prefectures are central-and-western while the 7 provinces are eastern, **the two sets of figures must not be merged. Maturity: deployed (policy rollout + pilot-school network).**

Case 6 | The Large-Scale Expansion of University Embodied-Intelligence / Robotics Programs — the hardest indicator on the discipline side [deployed]

This is the hardest indicator of "university embodied entry." In its 2026 revised undergraduate-major catalog, the Ministry of Education listed **"embodied intelligence" in the catalog for the first time,** supporting **9 universities** — Beihang, Beijing Institute of Technology, Beijing University of Posts and Telecommunications, Northeastern University, Harbin Institute of Technology, Nanjing University of Aeronautics and Astronautics, Shanghai Jiao Tong, Xi'an Jiao Tong, and Zhejiang University — in adding the program; **343 undergraduate institutions** nationwide already offer a robotics-engineering major; and **102 vocational colleges** have filed an intelligent-robotics-technology major (admitting students in 2025) [175] [176] [177]. These authoritative counts from the official major catalog and the Sunshine College-Entrance platform show that embodied/robotics education at the university level is **an accomplished fact at the level of discipline-building** — a deployment entirely different in nature from "humanoids as K-12 teachers": what enters is university programs and labs, not K-12 classroom teaching. **Maturity: deployed (discipline-building / admissions).**

Case 7 | Beijing's Class-Hour Mandate for Primary/Secondary AI General Education — a hard constraint on the teaching side [deployed]

Beijing's *Work Plan for Advancing Artificial-Intelligence Education in Primary and Secondary Schools (2025–2027)* mandates that, **from the autumn 2025 semester, all primary and secondary schools in the city offer AI general-education courses, no fewer than 8 class hours per academic year, covering the full span from primary to senior secondary;** a teacher "Hundred-Thousand Seed Program" is implemented in parallel [22] [164]. This is a formal local-government plan, and the **class hours are a hard constraint.** The "banned in examinations, advanced in teaching" boundary noted in §9.1 applies: Beijing's plan pushes hard on class hours on the teaching side, while the contemporaneous usage guidelines ban generative AI from answering questions directly on the examination side. **Maturity: deployed (mandatory coverage from autumn 2025).**

*Case 8 | Shanghai's Compulsory *Fundamentals of Artificial Intelligence* — hard coverage by grade [deployed]*

From autumn 2024, Shanghai has offered the local course ***Fundamentals of Artificial Intelligence** *in primary grade 4 and junior-secondary grade 7,** integrated into the regular teaching schedule (media reports cite 1 class hour per week, no fewer than 30 class hours per academic year) [23]. **Tiered scope:** the grade-4/grade-7 offering is from an official briefing (hard); the "30 hours / 1 hour per week" is from secondary edtech reporting (a medium-strength source — before citing the hard figure, the Shanghai Municipal Education Commission's original should be re-checked ``). **Maturity: deployed (from autumn 2024).**

Case 9 | Yuhang District, Hangzhou — "100 Schools + 2 Bases" full-system rollout — a district-county exemplar [deployed (context)]

UBTech, together with Yuhang District in Hangzhou, rolled out an AI-education project in which **100 primary and secondary schools plus 2 major bases** formally began instruction — a sample of district-county-wide full-system scaling [178]. An honest flag is required: this is a **single media source and dates relatively early (2020),** and the blue book uses it as a narrative exemplar of "district-county full-system" rollout and as historical depth, not as the latest scale. **Maturity: deployed (district-county level), early in time.**

Case 10 | Vocational Robotics Training and Industry–Education Integration — the vocational-education landing [deployed]

102 vocational colleges have filed an intelligent-robotics-technology major (admitting students in 2025); vocational colleges have built "national virtual-simulation training bases," realizing "theory–virtual–real integration" (disassembly/assembly, electrical control, virtual simulation, line commissioning, and programming training); and, per a planning figure, the talent gap in the robotics field is projected to reach **4.5 million** by 2025 [177] [179]. The 4.5 million must be flagged as a planning projection. **Maturity: deployed (vocational training / discipline-building).** The genuine training landing on the vocational side (such as quadrupeds entering vocational colleges for algorithm training — see Chapter 10 §10.7, the Hangzhou Vocational and Technical College case) is one of the few closure points of the capability–deployment gap.

Summary of the Chinese cases. Across the 10 cases, the hardest "deployed" evidence comes from the policy category (the class-hour mandate, base schools, discipline expansion) and from small robots entering primary and secondary schools; humanoids (Tien Kung Walker, Unitree G1) enter, in every instance, university research platforms. **Not a single case supports "full-scale humanoids deployed at scale into K-12 classrooms as teachers."**

9.3 International Cases (9): Software Agents at Scale, Social Robots Mostly in Pilots, and One Bankruptcy Warning

The picture of the international cases is equally clear, and it supplies two signals the Chinese cases rarely touch: **the genuine scaling of screen-based software AI tutors,** and **third-party empirical evidence that social-robot embodiments "mostly remain in pilots"** — the latter being the cleanest anchor for the capability–deployment gap.

Case 11 | Khanmigo (Khan Academy's AI Tutor) — software agent at scale [deployed at scale (software, not an embodiment)]

Khanmigo is the **only software AI tutoring agent deployed at scale** in education, but it is screen-based software, **not a robot embodiment.** K-12 student users rose from about 40,000 in 2023–24 to about 700,000 in 2024–25, while total users (teachers plus students) reached **about 1.4 million by April 2025;** roughly **350 U.S. school districts** use it daily; and New Hampshire's free deployment covers **50 districts, nearly 5,000 teachers, and 40,000 students** [11, 85, 180, 181].

The key capability–evidence gap: **as of the end of 2025, Khanmigo had still not completed a gold-standard RCT** (a matter of cost) [180]. The user/district counts are a vendor/annual-report figure (registered, not active), while New Hampshire's 50 districts / 5,000 teachers / 40,000 students is a state-government figure (hard). Khanmigo proves "the LLM brain's capacity to scale teaching," not "the teaching efficacy of an embodied robot" — a distinction that becomes, in Chapter 10 §10.7, the key basis for "the strong RCT evidence comes from text tutors and cannot be extrapolated to embodiments."

Case 12 | Amira Learning (an AI Reading Tutor) — state/national-level deployment [deployed at state level + national-level pilot]

Amira is an AI reading tutor, software rather than an embodiment. Vendor figures cite roughly 4–5 million students, 2,000–4,000+ districts, all 50 U.S. states, and about 19 countries (**a vendor range — use with caution**). Verifiable government-level deployments include: **Iowa** (in 2024, free statewide to all public and private elementary schools), **North Dakota** (statewide direct procurement for K–5), and the **Newark Public Schools** (adopted district-wide as a universal K–3 screener after a six-school pilot);

internationally, **Jordan's Ministry of Education** launched a national literacy program (a grade-4/grade-7 pilot with Arabic localization, transitioning to nationwide rollout after 12 weeks, with "up to 800,000 students" eligible to be enrolled) [86] [182] [87] [183]. The state-level deployments and Jordan are government-confirmed figures (strong); Jordan's "800,000" is an eligibility ceiling, not enrollment, and must be stated cautiously. **Maturity: deployed at state level (U.S.) + national-level pilot → rollout (Jordan).**

Case 13 | Google Gemini for Education / LearnLM — scale reach + small-sample RCT [deployed at scale + pilot RCT]

In 2025, Gemini for Education reached **10 million students across 1,000+ U.S. institutions** (a vendor reach figure). The efficacy signal comes from an **exploratory RCT** in UK classrooms in late 2025 (LearnLM + Eedi): **165 students across 5 UK secondary schools,** with LearnLM-supported students about 5.5 percentage points more likely to solve a new problem, and supervising teachers adopting **76.4% of its drafted messages with zero or minimal edits** [88] [95]. This must be flagged strictly: the RCT is a small-sample exploratory study (the abstract does not report significance), positioned as "promising but preliminary"; and it too is screen-based software, **not a robot embodiment.** The full evidence discussion belongs to Chapter 10 §10.7. **Maturity: deployed at scale (Gemini) + small-sample pilot RCT (LearnLM).**

Case 14 | NAO / Pepper + Aldebaran's Bankruptcy — a sustainability warning for a deployed embodiment [deployed but parent company in liquidation]

This is the most cautionary case in the chapter. NAO / Pepper are the most common social-robot embodiments in classrooms worldwide: Aldebaran (SoftBank Robotics) historically sold **roughly 20,000 NAO units plus 17,000 Pepper units** (across about 70 countries), with more than 17,000 units cumulatively deployed in the education market (a vendor historical-cumulative figure, not an active in-use count). But **the parent company, Aldebaran, entered bankruptcy liquidation in February 2025** (with cumulative debts of about €150 million), and the **IP and assets of NAO / Pepper were acquired at auction by Shenzhen Mornx (Maxvision),** which set up a French research-and-service branch [184] [185] [186].

The meaning of this **major industry event** reaches well beyond the rise and fall of one company: it shows that even a benchmark model "deployed" in tens of thousands of classrooms has a long-term sustainability that hinges on the parent company's commercial survival and supply-chain stability. **"Deployed" does not equal "sustainable"** — when a founding vendor such as Aldebaran heads into liquidation and its assets are acquired across borders, a school's long-term dependence on a single vendor's model is exposed as a structural risk. This warning, together with Case 18 (Moxie's shutdown), forms this chapter's twin reminder on "embodiment sustainability." **Maturity: deployed but parent company in liquidation → sustainability risk.**

Case 15 | A Global Third-Party Census of Social Robots — the best anchor for "deployed vs. piloted" [mixed: most remain in pilots]

This is **the cleanest third-party signal** for gauging the true maturity of social-robot embodiments. A peer-reviewed census recorded **206 real deployment cases across 28 countries and 43 models:** by country, the U.S. led with 74 and China with 16; by setting, special education / occupational therapy 65, primary/secondary schools 65, universities 39, and early-childhood 23; NAO and Pepper had the highest deployment counts [149]. Its **core conclusion is that most robots remain in the testing stage** (i.e., mostly pilots or short-term studies rather than mature, routine integration). The study counts "cases," not "units," and is the most powerful third-party corroboration of this blue book's "the capability–deployment gap is not yet closed" judgment: **the mainstream state of social-robot embodiments in real educational settings remains the pilot, not routine deployment at scale.**

Case 16 | The L2TOR Large-Scale Second-Language Social-Robot Study — a key near-zero-gain counter-example [research/pilot]

L2TOR is the largest study of its kind on robot-assisted second-language learning to date, and one of the most important **counter-examples** on the deployment side: 194 Dutch children, about 5 years old, were taught 34 English words by a NAO robot over 7 lessons, and **the robot-plus-tablet condition was not significantly better than the tablet alone,** with iconic gestures conferring no significant gain [150] [187]. This was an EU H2020 project (2016–2018, multi-site, peer-reviewed). From the standpoint of deployment efficacy it confirms Chapter 8's positioning of the language-learning track as "affective value clear, cognitive efficacy cautious"; the evidence details appear in Chapter 10 §10.3. **Maturity: research/pilot (project ended 2018).**

Case 17 | QTrobot (LuxAI) Special Education + a 69-Family Longitudinal Study — small-scale deployment + new research [deployed (small-scale) + pilot]

QTrobot (LuxAI) is used in special-education schools across several countries (named cases include the UK's Rodney House School and Bridge College; UNESCO cites its inclusive-education use); in late 2025, LuxAI, together with the Luxembourg Institute of Health and the University of Birmingham, launched **the first large-scale at-home longitudinal study: 69 families** receiving robot-assisted early-development support for children with autism [188]. The "69 families" is a precise third-party-institution figure, while "used in schools across several countries" is a vendor figure with no hard count. The evidence firewall from Chapters 8 and 10 carries over — **clinic efficacy in special education / ASD does not equal universal classroom efficacy** — and this longitudinal study is precisely an effort to fill the long-term evidence gap for the at-home setting. **Maturity: special-education deployed (small-scale) + new longitudinal study (pilot).**

Case 18 | Embodied "Moxie" Shutdown — a sustainability warning for cloud-dependent companion robots [shut down]

Embodied shut down in late 2024 after a failed funding round, and Moxie — an US$800 cloud-dependent children's companion robot — was "bricked" in early 2025 when its servers went dark, with no refunds, the company later attempting to open-source it [189] [190] (public sales figures are unknown; no number is cited). This is a **landmark warning on the sustainability of LLM/cloud-dependent companion robots:** when the core intelligence depends on a vendor's cloud service, the embodiments already sold are instantly disabled the moment the company shuts down. Together with Case 14 (Aldebaran), it forms this chapter's twin warning on "embodiment sustainability." **Maturity: shut down.**

Case 19 | ABii / Van Robotics (U.S. K-5 Tutoring Robot) — small-scale deployment + independent research [deployed (small-scale) + independent field study]

ABii is used by "thousands of children" across 30+ U.S. states and in the UK, Germany, the UAE, and Qatar; in 2025 there was an **independent, long-term in-school field study** (peer-reviewed) [191]. An honest flag is required: the "thousands / list of states" and the vendor's early effect claims (67% / 34%) are vendor figures and should be used with caution; it is the 2025 independent study that is the basis for its maturity assessment. **Maturity: deployed (small-scale) + independent field study.**

Summary of the international cases. What is "deployed at scale" are **screen-based software AI tutors** such as Khanmigo / Amira / Gemini; social-robot embodiments (NAO/Pepper/QTrobot/ABii) are mostly small-scale deployments or pilots, and a third-party census states plainly that "most remain in the testing stage"; and two sustainability warnings have surfaced — Aldebaran's bankruptcy and Moxie's shutdown. **This is entirely consistent with the Chinese cases: routine deployment at scale on the embodiment side has, in 2026, not yet arrived.**

9.4 Western Humanoids: Industrial Only, No Educational Deployment

Turning the gaze from "embodiments that have landed" to "the Western humanoids repeatedly juxtaposed with the future of education" yields this chapter's most clear-cut, least ambiguous judgment: **Western humanoid-robot vendors barely enter education; they enter only industry.**

The real deployment of representative Western humanoid vendors — Figure, Tesla Optimus, 1X NEO, Apptronik and the like — is concentrated on the **industrial production line:** Figure 02 piloted on the BMW Spartanburg line, Tesla Optimus at the Fremont factory, and related robots tested on lines such as Mercedes [143] [8]. In education, these vendors have **no traceable deployment.** Even Figure's widely circulated Helix "doing chores" footage belongs to the vendor-release-video / early-pilot category already flagged in Chapter 4, not to any real deployment in an educational setting.

This reality carries an important narrative implication: when the international media set "humanoid robots" alongside "the future of education," the blue book must point out coolly that — **in the**

deployment map of Western humanoids, the education cell is essentially blank; they are the next generation of industrial-automation equipment, not teachers in the classroom.

Seen more deeply, the divergence between Western and Chinese humanoids in education reveals two entirely different routes "into education." Western vendors (Figure / Tesla / 1X / Apptronik) take an "industry-first" route, positioning the humanoid as a replacement for or complement to line workers, with education never at the core of their commercial map; the Chinese humanoids that can enter education (Unitree, UBTech's Tien Kung Walker) take a "research-platform" route, supplying the humanoid to university labs as a research tool. The common ground of the two routes is precisely that — **neither is "the humanoid robot as a primary- or secondary-school teacher."** East or West, "full-scale humanoids entering K-12 classrooms at scale" remains, in 2026, a concept and a demonstration rather than a deployed reality. This is the empirical basis of the blue book's "humanoid-in-education, cooled" judgment.

9.5 Channel Build-Out Abroad: The Global Presence of Dobot and Makeblock

In contrast to the "heavy hardware, light educational deployment" of the humanoid track, China's coding/STEM robot vendors show a **genuine and verifiable scale channel build-out** in educational exports — consistent with the "already at commercial scale" positioning of the STEM-kit track in Chapter 8.

Educational-entry status [deployed (channel at scale)]. Dobot (Yuejiang Technology) has entered **6,500+ institutions across 80+ countries;** Makeblock covers **140+ countries with more than 20 million users** [192]. The research verified these figures as counts of institutions and countries in the Chinese original (the units *jia* / *guo*, "institutions/countries"), carrying **no currency-misreading risk** — unlike the *yi yuan*-versus-billion problem that Chapter 13 works hard to correct, the "6,500+" and "20 million+" here are counts of institutions and users, confirmed beyond doubt.

Two **single-source, pending-verification** amounts must be presented honestly side by side (PENDING-secondary): media reports state that VEX received roughly **US$18 million** in Career and Technical Education (CTE) funding (covering 12 states and 5,000 sets of equipment) and that Makeblock raised roughly **US$25 million** in a Series D [192] — these amounts are from a single, uncross-checked media source, which the blue book flags as pending, making no statement of them as accomplished fact.

Handling "verified deployment counts" and "pending amounts" separately is an expression of research honesty: within the same source, a verified institution count does not certify that an amount is also verified. The full capital-side analysis of financing belongs to Chapter 13.

This channel scale confirms the judgment of Chapter 8: in STEM/coding, the most mature track, Chinese vendors are deployed at scale domestically and, through their global channel build-out, abroad, forming a considerable presence — one of the most solidly evidenced "deployed" maps in the educational-robot industry.

9.6 Chapter Summary

Using the dual yardstick of educational-entry status and evidence tier, this chapter has mapped 19 real-world deployment cases (Fig. 11), which distil into five judgments.

The first concerns who actually reaches the classroom at scale: **software agents and small robots, not humanoids.** Screen-based software AI tutors (Khanmigo's 1.4 million total users (teachers and students; about 700,000 K-12 students) [11], Amira's state/national-level deployment, Gemini's 10-million reach [88]) and UBTech's Alpha Mini-class small coding robots [163] are the most solidly evidenced "deployed" samples.

Where a humanoid does enter education, it enters as **a university research platform.** Unitree's G1 (roughly three-quarters of revenue from research and education [107]) and UBTech's Tien Kung Walker [167] land in the university research layer, not the primary/secondary teaching layer; the "300+ units" is a vendor claim and must be tracked separately from "already delivered into universities."

The hardest landing of all comes from the policy and discipline side. Beijing's ≥8 class hours per academic year [22], Shanghai's compulsory grade-4/grade-7 course [23], the 509 base schools [173], and the 9 universities first offering an embodied-intelligence major plus 343 offering robotics engineering [175] are the hardest indicators of "deployed" embodied/AI education across universities and primary/secondary schools — with the "banned in examinations, advanced in teaching" boundary held throughout [164].

Social-robot embodiments, by contrast, are mostly in pilots, and carry sustainability warnings. A third-party census reports that, of 206 deployment cases, "most remain in the testing stage" [149]; L2TOR's 194 children showed near-zero gain [150]; Aldebaran went bankrupt in February 2025, its assets acquired by Shenzhen Mornx [184] [185], and Moxie shut down [189] — "deployed" does not equal "sustainable."

Western humanoid educational deployment, finally, is blank. Figure / Tesla / 1X / Apptronik have only industrial deployment [143], and any "humanoid teachers in the classroom" claim is prohibited.

A look ahead. Chapter 9 has answered who genuinely entered education and how hard the evidence is; **Chapter 10** answers the deeper question — for these deployed or piloted robots, **what are the learning outcomes?** Chapter 10 takes up the "strength of evidence" flags from each track in Chapter 8 and the deployment cases of this chapter, developing effect sizes scenario by scenario, with a strength-of-evidence color scale and methodological cautions, and recording null and adverse effects with equal prominence — above all, strictly distinguishing the "clinic/research-effective ≠ universally classroom-effective" theme this chapter has repeatedly stressed, and using real landings such as Hangzhou Vocational and Technical College's quadruped training as evidentiary corroboration for a "closure point" of the capability–deployment gap. Vendor financials, shipment rankings, and the capital "ice and fire" belong to Chapter 13.

Chapter 10 Learning Outcomes and Evidence: Scenario-by-Scenario Strength Grading

> *Learning Outcomes and Evidence — Locally Robust, Globally Cautious: Scenario-by-Scenario Strength Grading, with Null and Adverse Effects Equally Reported*

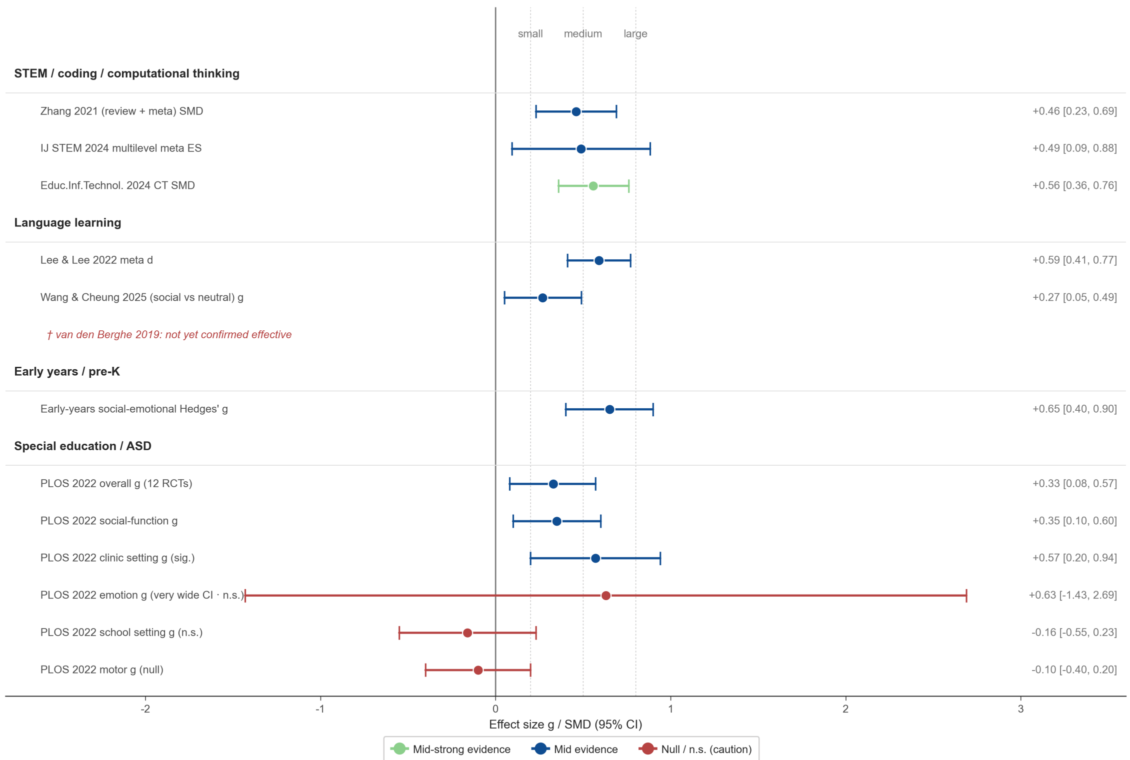
10.1 The Iron Rules of Evidence and the Methodological Cautions

Chapter 8 gave the market scope of the five sub-tracks, and Chapter 9 mapped the vendor-deployment landscape — but one more fundamental question has remained unresolved throughout: **do educational robots actually help students learn better?** This chapter is the blue book's central evidence chapter, and the last line of defense against the overstatements of "treating a vendor demo as accomplished fact" and "treating a rise in engagement as a learning gain."

For each scenario, this chapter strictly enforces six evidence rules:

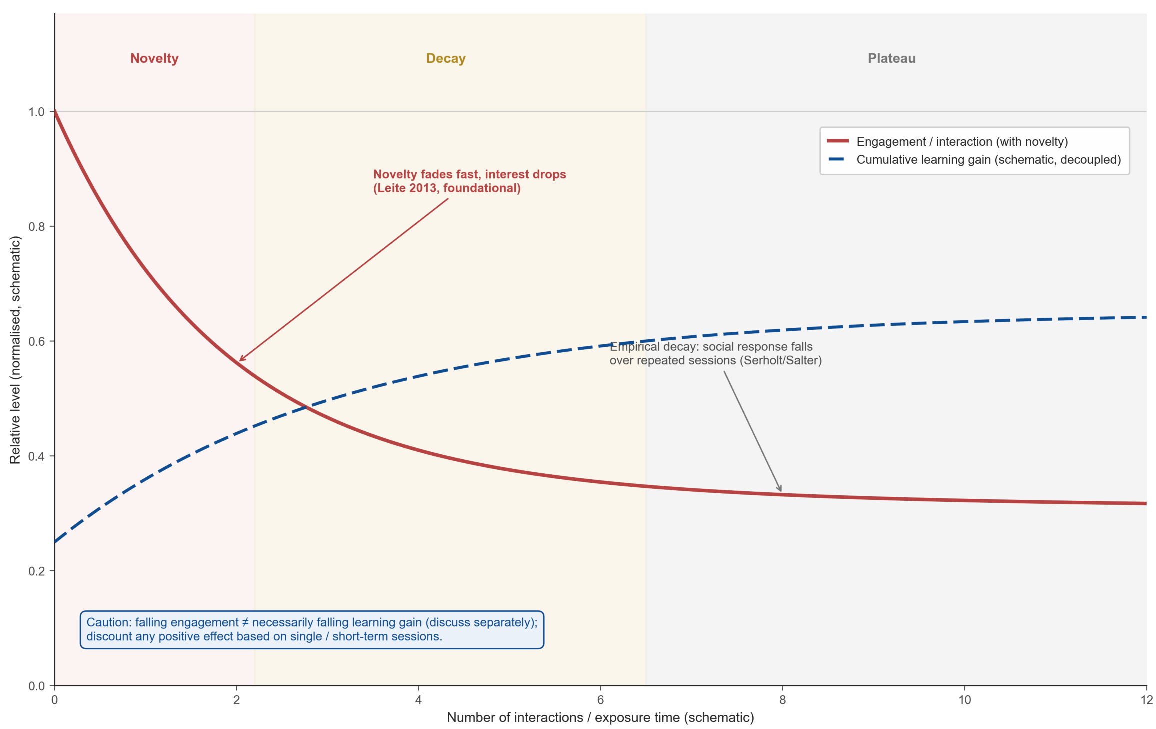
1. **No fabricated effect sizes.** All *g*/SMD/*d* values are taken only from meta-analyses and empirical studies verified in the research; where no verified source exists, a qualitative statement is preferred. 2. **Null and adverse effects are reported with equal prominence.** Reverse signals — the null effect on the motor dimension, non-significance in school settings, LLMs reducing engagement — are presented at the same length as positive effects. 3. **Strength of evidence (strong/medium/weak) and limitations are flagged by scenario.** 4. **The novelty effect is discounted.** Any positive effect based on a single session or short term is discounted — the most robust methodological caution in this chapter (see §10.6). 5. **Affective gains are distinguished from cognitive gains.** The former often masks the latter; the most common positive effects of robots cluster on the affective/motivational dimension. 6. **Beware the risk of extrapolation across settings: clinic/research-effective ≠ universally classroom-effective** — the core trap of the ASD track (§10.5).

The systematic methodological threats running through the chapter must be stated first: the **novelty effect** (pervasive in scenarios 2/3/5/6), **publication bias** (under-reported in most meta-analyses), **low study quality** (only 17.5% rated "strong" in ASD [161]), a **high share of quasi-experiments**, **small-sample, single-session dominance**, and **affective gains ≠ cognitive gains.** The by-scenario effect-size forest plot appears in , and the novelty-effect decay curve in .



Sources: Zhang 2021, IJ STEM 2024, Educ.Inf.Technol. 2024, Lee & Lee 2022, Wang & Cheung 2025, PLOS One 2022 (ASD), early-years social-emotional meta 2025. Caution: many are quasi-experimental / single-session, discount for novelty; clinic-effective ≠ universally classroom-effective (school g=−0.16, n.s.).

Figure 12.



Sources: Leite et al. (2013, Int. J. Social Robotics), foundational on novelty effect; Serholt/Salter empirical-decay review (2024). Curves are schematic, not measured data; they illustrate the robust caution that short-session positive effects are unreliable.

Figure 13.

Cross-scenario methodological cautions (essential reading for drafters and readers). The six systematic threats warrant full explanation before the scenario-by-scenario evidence, because they determine that the great majority of effect sizes in this chapter should be read cautiously:

- **Novelty effect:** the excitement and engagement children feel when faced with a novel robot drops sharply after a few interactions. It systematically threatens scenario 2 (language), scenario 3 (early childhood), scenario 5 (social-emotional), and scenario 6 (LLMs), so that any positive effect based on a single or short-term session may be a product of "novelty" rather than "learning gain." This is the most robust and most binding basis for discounting in this chapter (§10.6 treats it specifically). - **Publication bias:** positive results are more likely to be published, leading meta-analyses to overestimate true effects. Most meta-analyses in this field under-report bias tests; the only one to clearly report Egger's test and find **no significant bias** is the CT meta-analysis [155] — which is an important reason the STEM-scenario evidence is relatively credible. - **Low study quality:** in the ASD track, for example, only **17.5% of included studies were rated "strong" for quality** in the PLOS One (2022) meta-analysis [161], with many studies of low-quality design, weakening the generalizability of the conclusions. - **High share of quasi-experiments:** many studies use quasi-experimental rather than strictly randomized controlled (RCT) designs, and cannot fully rule out selection bias and confounding. - **Small-sample, single-session dominance:** most language-learning studies have fewer than 80 participants and are single-session [193]; the median follow-up in the ASD track is only about 10 weeks — short windows and small samples struggle to capture long-term learning effects. - **Affective gains ≠ cognitive gains:** the most common, most robust positive effects of robots cluster on the affective/motivational dimension (reduced anxiety, increased engagement), while cognitive-level learning gains are often uncertain. Mistaking the former for the latter is the most pervasive substitution in efficacy claims for educational robots.

These six threats are not an abstract methodological checklist but the concrete basis for "discounting" each scenario's effect sizes — behind every *g* value, one must ask which design it came from, whether it has passed through novelty-effect decay, and whether it measures affect or cognition.

10.2 Scenario One — STEM / Coding / Computational Thinking (evidence strength: medium–strong)

STEM/coding/computational thinking (CT) is the **most robustly evidenced** of all scenarios, and the track this blue book can state relatively positively.

Core effect-size evidence (E1): several independent meta-analyses converge on a **medium positive effect (*g*/SMD ≈ 0.46–0.56)** —

- Zhang et al. (2021), a review plus meta-analysis: **SMD = 0.46 (CI 0.23–0.69)** [194]; - *IJ STEM Education* (2024), a multilevel meta-analysis (21 studies / 30 effect sizes): **ES = 0.488 (CI 0.094–0.882, *p* < .05),** with the CI's lower bound close to 0.09 [195]; - *Educ. Inf. Technol.* (2024), a CT meta-analysis (27 experimental/quasi-experimental studies): **SMD = 0.558,** with the strongest effect in early childhood, and **Egger's test finding no marked publication bias** [155].

The earlier Benitti (2012) systematic review offers qualitative support but lacks a quantitative assessment [196]; an early-childhood CT meta-analysis (PMC, 2024) is consistently positive in direction but uses non-uniform measurement instruments [197].

These three independent 2021–2024 meta-analyses converge tightly on effect size (0.46 / 0.488 / 0.558) — consistent in direction and close in magnitude — and this cross-study, cross-year consistency is itself a mark of robust evidence: it lowers the risk of bias from "a single team, a single definition." More remarkably, the CT meta-analysis's reported Egger's test found **no marked publication bias** [155], meaning this medium positive effect is unlikely to be an artifact of "selective publication of positive results."

A cautionary note: relatively robust as the evidence is, three limitations qualify it. Most studies include quasi-experiments rather than pure RCTs and cannot fully rule out selection bias; STEM learning attitudes are mostly self-reported and skew optimistic; and some meta-analyses include few studies ($k \approx 21$ [195]), with the IJ STEM meta-analysis's CI lower bound already approaching 0.09, indicating that, though significant, the effect is not thickly bounded at the bottom. **The conclusion to adopt:** the medium positive effect for STEM/computational thinking (g/SMD ≈ 0.46–0.56) can serve as the blue book's relatively positive evidentiary pillar, with at least one CT meta-analysis finding no publication bias [155] and the strongest effect in early childhood — which also forms a two-way industry-and-evidence corroboration with Chapter 8's market judgment that "STEM kits are the only track already at commercial scale."

10.3 Scenario Two — Language Learning (evidence strength: medium, cautious)

Language learning is the scenario most strictly constrained by **evidence firewall E2:** one must **not claim "proven effective," and affective gains > cognitive gains.**

A wide effect-size span (E2):

- Lee & Lee (2022), a meta-analysis of robot-assisted language learning: **d = 0.59 (SE 0.09)** [198]; - Wang & Cheung (2025): social behaviors vs. a neutral condition, **g = 0.269,** but mostly K-12 English vocabulary, samples < 80, single-session [193]; - a children's language-development meta-analysis (2003–2023, 27 empirical studies / 70 effect sizes / 1,544 participants): a medium-to-large positive effect, **most significant on the affective dimension** [199]; - Randall (2019): a qualitative review of 79 studies, with mixed evidence [157]; **- the L2TOR large-scale counter-example (near-zero gain):** 194 Dutch children, about 5 years old, taught 34 English words by a NAO, with **the robot-plus-tablet condition not significantly better than the tablet alone** (the largest study of its kind, 2016–2018, multi-site, peer-reviewed) [150] [187].

The effect size spans widely from $g \approx 0.27$ to $d \approx 0.59$, **and cognitive gains are often masked by affective/motivational gains;** the great majority of studies are single-session/small-sample, and the novelty effect is hard to rule out. Most striking of all, **the largest study of its kind to date, L2TOR, obtained precisely "near-zero gain"** — robot-plus-tablet not significantly better than tablet alone [150] — a counter-example as important as every positive effect size in this scenario.

The conservative counterpoint that must be presented alongside (E2 red line): the authoritative review by van den Berghe et al. (2019) offers a clear, cautious judgment — **"robots cannot yet be confirmed as effective language tutors"** [158]. This conclusion carries great weight because it comes from a widely cited systematic review of the field, and it rests on two concrete, pointed negative observations. In preschool second-language (L2) learning, student engagement **begins to decline within three sessions** — a direct manifestation of novelty-effect decay in the language scenario. And when the robot exhibits gaze behavior, **grammar learning is actually worse**, suggesting that certain anthropomorphic social cues from the robot may distract children from the language content itself.

These two observations together form a forceful rebuttal to the intuition that "rising engagement means effective learning": students facing a robot are less anxious and more engaged (Chapter 8 §8.3), but this does not automatically translate into better language acquisition. **In the language-learning scenario, this blue book always presents this conservative conclusion alongside the data; language-learning efficacy is not claimed to be "proven effective," and the presentation of affective/motivational gains must be set against the uncertainty of cognitive gains.** The affective value of the robot as a "low-anxiety language-practice partner" is clear, but its cognitive efficacy as an "effective language tutor" has not been confirmed — the two must be stated separately.

10.4 Scenario Three — Early Childhood (evidence strength: medium)

The core feature of the early-childhood scenario is that **affective gains exceed cognitive gains, and the novelty-effect risk is highest.**

Effect-size evidence (E4): early social-robot interventions (covering both ASD and typically developing, TD, children) had a **Hedges' g = 0.651** effect on social-emotional development [200]; the children's language-development meta-analysis further shows that **younger children gain more affectively and older children gain more cognitively** [199]. Combined with scenario 1, the computational-thinking effect is also strongest in early childhood [155].

A cautionary note (discounting the strength): $g = 0.651$ looks considerable, but it calls for strict discounting on two grounds. Measuring young children's cognitive constructs is itself highly contested: young children cannot reliably complete standardized cognitive tests, and researchers often fall back on behavioral observation or proxy indicators, leaving construct validity in doubt. And young children are especially sensitive to novel stimuli — the group most affected by the novelty effect of any age range — so that **a short-term rise in engagement ≠ a long-term learning gain:** the appeal of a talking, moving robot to a young child is easily misread as educational efficacy.

The evidence shows a consistent age gradient: **younger children gain more affectively, older children gain more cognitively** [199]. This implies that the true value of robots in early childhood lies more in affective companionship and the priming of social-emotional skills than in accelerating cognitive/subject learning. Early childhood is the track with the widest gap between market enthusiasm and evidence strength (Chapter 8 §8.4 shows it accounts for only about 1% within a strict statistical scope), and this blue book holds to the positioning of "affective gains clear, cognitive gains cautious, novelty effect must be discounted" — which is also the evidentiary basis for Chapter 14's discussion of governing young children's emotional dependence.

10.5 Scenario Four — Special Education / ASD (evidence strength: medium; extrapolation in doubt)

ASD intervention is the **only track with relatively abundant RCT meta-analytic support,** and the most critical scenario for **evidence firewall E3: clinic-effective ≠ universally classroom/home-effective.**

Core effect-size evidence (E3, the most robust single item) — the PLOS One (2022) meta-analysis reviewed 40 studies (including 17 RCTs) and meta-analyzed 12 RCTs / 346 participants [161]:

- overall Hedges' g = 0.33 (CI 0.08–0.57); - social functioning g = 0.35, significant — the most reliable positive effect of ASD intervention; **- emotion dimension g = 0.63 but not significant** (with an extremely wide CI of −1.43 to 2.69); **- motor dimension g = −0.10, a null effect.**

Null/adverse effects and setting moderation (E3 red line, must be reported with equal prominence) — the same meta-analysis's setting-moderation analysis reveals the most critical extrapolation risk:

- clinic setting g = 0.57, significant; - home setting g = 0.16, not significant; - school setting g = −0.16, not significant [161].

That is, the significant positive effects of robot ASD intervention come mainly from the **controlled clinic environment,** and once extrapolated to the school classroom or the home, the effect becomes non-significant (school g = −0.16, home g = 0.16, both non-significant) or even turns negative in direction. This setting-moderation result is one of the most important — and most easily buried — findings of the entire evidence chapter: it means the "strongest evidence" halo of the ASD track in fact depends heavily on the clinic, a controlled, one-on-one environment led by professional therapists, whereas the school classroom where real educational deployment occurs is precisely the setting in which the effect vanishes. Layered on top are the study-quality limitations — **only 17.5% of studies rated "strong" for quality** [161] — and a median follow-up of only about 10 weeks, where the short window likewise struggles to rule out the novelty effect. Funnel-plot analysis showed no marked asymmetry, suggesting that the publication-bias risk of this particular meta-analysis is relatively controlled.

Supplementary evidence and literature concentration: Zhong (2025), a three-level meta-analysis, agrees in direction with the above, holding that robots have a positive effect on ASD social-emotional skills [201] (source fields pending verification, see Appendix F). At the platform level, NAO is the de facto standard model of this track, appearing in about 39% of ASD-related robot studies [53] [202] [203] [162] — a concentration that both eases cross-study comparison and raises a concern: the evidence base depends heavily on a single model, and whether its conclusions generalize to other social-robot embodiments is itself in doubt.

The conclusion to adopt (a strong track, but cautious): robots have a small-to-medium significant effect on ASD **social functioning** ($g \approx 0.33–0.35$), with younger children gaining more; but **one must never claim "clinic-effective" to be "universally classroom/home-effective"** — the emotion and motor dimensions are null or non-significant, and the home and school settings are both non-significant. This is the most easily abused trap within the strongest-evidence track.

10.6 Scenario Five — Social-Emotional Effects and Novelty-Effect Decay (evidence strength: medium, cautionary)

Novelty-effect decay is the **most robust cautionary finding** in the educational-robot field, and the systematic basis for discounting all the "short-term positive effects" above (**evidence firewall E5**).

Core evidence (E5):

- Leite et al. (2013): the foundational definition of the novelty effect — **the novelty effect fades quickly, after which interest declines; short-term sessions cannot reliably predict long-term effects** [204]; - empirical decay (Serholt/Salter): social responsiveness declines over repeated interactions, and boredom rises after long-term exposure [205]; - a long-term study of an adaptive robot tutor in higher education [206]; a best-evidence synthesis on empathy development showing mixed evidence [207].

The mechanism of novelty-effect decay can be summarized as a decay curve (Fig. 13): engagement is high early in the interaction, declines monotonically as interactions accumulate, and finally levels off at a plateau far below the initial level. Leite et al. (2013) first systematically defined this phenomenon and, on that basis, advanced a claim devastating for educational evaluation — **short-term sessions cannot reliably predict long-term effects** [204]. Serholt and Salter's long-term observations provide empirical support: social responsiveness declines over repeated interactions, and children's boredom rises after long-term exposure [205]. In other words, the conclusion of a single-session study reporting that "the robot significantly improved engagement" may completely reverse weeks later.

The conclusion to adopt: any positive effect based on a single/short-term session should be discounted (Fig. 13). One subtlety must be noted at the same time — **a decline in engagement ≠ a necessary decline in learning gains,** and vice versa: van den Berghe's (2019) counter-example shows that when the robot gazes, engagement may rise while grammar learning is actually worse [158]. Engagement and learning gain are two constructs that must be measured and discussed separately, not simply equated. This also echoes Chapter 7's judgment that "engagement ≠ learning gain" — the high-engagement signals captured by multimodal systems cannot be read directly as better learning outcomes.

10.7 Scenario Six — New Evidence in the Foundation-Model Era (evidence strength: weak–medium)

The efficacy of educational robots in the LLM era is the scenario most strongly constrained by **evidence firewalls E6/E7: promising but early; the strong RCT evidence comes from text-based LLM tutors and cannot be extrapolated to embodied robots.**

Key evidence (E6/E7):

- **TalBot (ChatGPT × Furhat):** a pilot/feasibility study with only 13 preschool children, **reporting no controlled learning gain** — controlled evidence for embodied LLM robots with children is almost blank [208]; - **Tutor CoPilot:** a relatively strong **RCT** (>700 tutors / >1,000 students), with students' mastery probability **+4 percentage points,** and students taught by lower-rated tutors **+9 percentage points** — but this is a **text-based LLM tutor, not a robot** [209]; - **Google LearnLM (Eedi):** an RCT (165 UK students aged 13–15), with supervising teachers approving **76.4% with zero or minimal edits** (a process-quality proxy, not an attainment effect size); likewise text-based, not a robot [210] (source fields pending verification, see Appendix F); - **counter-evidence (E7):** a large-class quasi-experiment on GPT programming showed it **reduced overall engagement** while raising the attainment of adopters, **widening the divide** [211].

Why it cannot be extrapolated (the most critical logic of this scenario): there is an easily overlooked evidence mismatch here. The media and vendors often juxtapose exciting RCT numbers such as Tutor CoPilot's +4 pp / +9 pp and LearnLM's 76.4% with the "educational robot" promise narrative, creating the illusion that "LLM robots have been proven effective by RCT." But the fact is that all this strong evidence was generated in the **text/screen-based LLM-tutor** context, and its mechanism (real-time assistance to a human tutor, generating Socratic follow-up questions) is a different intervention from "an embodied robot accompanying a child in physical space" — the efficacy of the former cannot be mechanically transferred to the latter. This logic is entirely consistent with the judgment repeatedly stressed in Chapter 5: **Khanmigo is a screen-based software agent, not a robot embodiment,** and what it proves is "the LLM brain's teaching capacity," not "the teaching efficacy of an embodied robot." Once the brain is connected to a body and physical presence and multimodal social cues are introduced, whether efficacy holds, how the novelty effect operates, and how hallucination risk is amplified all lack controlled evidence.

The conclusion to adopt (banning any efficacy claim): (a) controlled evidence for embodied LLM robots with children is almost blank (only the TalBot pilot, 13 children, no control [208]); (b) the relatively strong RCT evidence (Tutor CoPilot, LearnLM) **comes entirely from text-based LLM tutors and cannot be extrapolated directly to embodied robots;** (c) reverse signals — reduced engagement, a widened divide — have already appeared [211]. **In this blue book, LLM robots are positioned as "promising but early," and any efficacy claim is banned.** This cautious positioning will be translated, in Chapter 14, into governance requirements for model hallucination, academic misdirection, and emotional dependence.

10.8 A Few Closure Points of the Capability–Deployment Gap: Vocational Quadruped Training (evidence strength: weak)

The six scenarios above paint a picture of "cautious efficacy." But the blue book's "capability–deployment gap" spine must also record its **few closure points** — settings that are both genuinely deployed and host observable learning activity. The most representative is **quadruped robots entering Chinese vocational colleges for perception and algorithm training.**

Deployment fact (deployed, vocational-college classroom training): Hangzhou Vocational and Technical College (HZPT) introduced the Unitree Go2 into the professional teaching of its School of Internet of Things Technology, where students use the lidar point clouds plus 4K imagery streamed back by the robot dog for spatiotemporal alignment, writing algorithms to identify campus-violation targets, and teachers say this lets students "grasp machine learning and computer vision more intuitively" [138]; the RoboUniversity certification system Unitree co-runs with Stanford and OpenMind also offers, for the Go2/G1 platform, courses from programming fundamentals to robot integration [127]. This is the real closure, on the vocational side, of the judgment Chapter 9 argued repeatedly: **the quadruped is the most realistic platform for embodied intelligence to enter education at present — its motion control is engineered, its price threshold is low (the education version is in the low-thousands-of-dollars range), and it is already used for real algorithm and perception training in vocational/university classrooms.**

But the evidence firewall does not relax here either. Two cautionary boundaries must be made clear: first, **"deployed" means teaching activity genuinely takes place, not that learning outcomes have been quantified** — the HZPT case is a deployment fact of engineering training, with **no controlled-design learning-gain data** at present, and one cannot claim from it that "quadruped robots improved learning outcomes"; second, **the landing is strictly limited to vocational/university settings,** a training scenario aimed at skills practice and algorithm work, and **cannot be extrapolated to "robot dogs have entered K-12 classrooms at scale."** In other words, this "closure point" closes the question of "whether a capability can be used in real teaching activity," not "whether efficacy has been confirmed by evidence" — the two must be stated separately. This is precisely the consistent extension of this chapter's methodology: the reality of deployment and the provability of efficacy are two different yardsticks.

The conclusion to adopt: vocational quadruped training is one of the few observable closure points of the "capability–deployment gap" and can serve as deployment corroboration that "embodied capabilities are entering real teaching activity" [138]; but its learning outcomes are not yet quantified, so it is **positioned as "deployment real, efficacy not quantified,"** stated on a separate track from the effect-size evidence of the scenarios above, and merged into no "proven effective" conclusion.

10.9 Three of the Most Robust Findings: The Core Evidentiary Pillars

Synthesizing the chapter's scenario-by-scenario evidence grading (Fig. 12), this blue book distills three findings that are **the most robust and can serve as core evidentiary pillars:**

1. **STEM / computational thinking: a medium positive effect** (g/SMD ≈ 0.46–0.56), with at least one CT meta-analysis finding no publication bias, strongest in early childhood [194] [155] [195]. This is the only track that can be stated relatively positively.
2. **ASD social functioning: a small-to-medium significant effect in RCT meta-analyses** (g ≈ 0.33–0.35), with younger children gaining more [161] — but **clinic-effective does not equal universally classroom/home-effective,** with the motor/emotion dimensions and the school/home settings all non-significant [201].
3. **Novelty-effect decay: a multi-source consistent caution** — engagement/responsiveness declines over time, and short-term effects are unreliable [204] [205]. This is the systematic basis for discounting all short-term positive effects.

Together these three pillars sketch a picture of "locally robust, globally cautious": the only track that can be stated relatively positively is STEM/computational thinking; the ASD social-functioning evidence, though relatively strong, is severely limited by the risk of extrapolation across settings; and the novelty effect is a yardstick of discount hanging over all short-term positive effects. By contrast, the cognitive efficacy of language learning, early childhood, and social companionship, as well as the overall efficacy of LLM robots, is at present insufficient to support a claim of "proven effective." This blue book would rather acknowledge the insufficiency of the evidence than use an optimistic single value to mask the true state of the evidence — which is precisely the empirical root of the seed of the Chapter 15 policy recommendation "sufficient validation, teacher-led, then deploy," and the evidentiary response to the lesson of South Korea's AI digital textbook "rapid entry into schools and backlash" (Chapter 14).

10.10 Chapter Summary

This chapter graded the evidence for the learning outcomes of educational robots strictly by scenario (Fig. 12, Fig. 13), and its core stance can be distilled as follows: **the evidentiary picture of educational robots is "locally robust, globally cautious" — STEM/computational thinking has a medium positive effect that can be stated positively, the ASD social-functioning evidence is relatively strong but extrapolation across settings is in doubt, language learning and LLM robots cannot be claimed proven effective, and the novelty effect is the most robust basis for discounting throughout; while a "deployment closure point" such as vocational quadruped training closes the question of capability entering real teaching activity, not of efficacy being confirmed.** Null and adverse effects (the null effect on the motor dimension, non-significance in school settings, LLMs widening the divide, L2TOR's near-zero second-language gain) have been reported with equal prominence alongside positive effects.

This chapter's evidentiary stance serves the blue book's "capability–deployment gap" spine directly: however dazzling a vendor's demonstration of a frontier capability, it must withstand evidence testing that is by scenario, by strength, and inclusive of negative signals; a rise in engagement/affective gains does not equal a cognitive learning gain.

A look ahead. Part III (Applications) is now complete — Chapter 8 gave the full track spectrum, Chapter 9 mapped the landing landscape, and Chapter 10 completed the evidence grading. **Part IV (Chapters 11–13)** turns to the industry view: Chapter 11 reconstructs the value chain (adding the VLA algorithm layer), Chapter 12 gives the master account of market and TAM with an honest look back at historical forecasts, and Chapter 13 presents the competitive landscape and the capital "ice and fire." This chapter's cautious evidentiary stance will also, in **Chapter 14** (Safety, Ethics, and Governance), be translated further into governance requirements for emotional dependence and academic misdirection, and converge in **Chapter 15** into the seed of the policy recommendation "sufficient validation, teacher-led, then deploy" (echoing the lesson of South Korea's AI-digital-textbook backlash).

Part IV Industry and Market

Value-chain reconstruction · market size and TAM · competitive landscape and capital

Turning to the industry view, this part traces the 2026 reconstruction of the value chain — a new model/embodied-AI algorithm layer and upstream bottlenecks that move forward to reducers and dexterous hands — gives market size in conservative, multi-source ranges, and portrays a capital split: a chill for traditional educational robots and a frenzy for embodiment and humanoids.

Chapter 11 Value-Chain Reconstruction: From the Hardware Chain to the VLA Layer and Upstream Bottlenecks

> ***Scope and de-duplication note.*** *This chapter addresses one thing only: the structural reconstruction of the industry chain — how the seven-layer baseline has evolved, which new layer has appeared, where the bottleneck links have moved, and how value is being redistributed along the chain. Market sizing and the total addressable market (TAM) belong to Chapter 12; vendor financials and shipment rankings belong to Chapter 13; the hardware parameters of the humanoid body (degrees of freedom, cost) belong to Chapter 6. Where those topics are referenced here, they are mapped structurally, not recomputed. The rationale for "adding a layer" shares its source with Chapter 3, and the technical account of VLA evolution continues from Chapter 4. > >* ***Conservative caveat.*** *The "seven layers + one new VLA layer" is a multi-source* ***2026 reconstruction judgment,*** *not a conclusion from the original white papers; the "foundation-model and embodied-AI service layer" is this report's proposed naming, flagged* ***[2026 extension].*** *Operations and maintenance accounting for 50–60% of total cost of ownership is from a single media source, used only qualitatively for cost structure; all local targets and base-school figures follow the policy original.*

11.1 Introduction: When the "Brain" Becomes a Layer of Its Own

When the 2019 edition of the *Global Educational Robots White Paper* drew the industry chain, the organizing logic of the whole chain was "centered on the body hardware": from the most upstream AI chip, through hardware manufacturing, system platforms, application services, content supply, and system integration, to the user-facing channel, seven links stacked in sequence ([2]). In that framework, "intelligence" was assumed to be embedded in the hardware and the platform and did not constitute an industry layer that needed to be depicted on its own — because at the time, a robot's "intelligence" was assembled from rules, sensing, and scripts, with no "brain" that could be priced, supplied, or iterated separately.

The 2026 industry picture has undergone a **structural reordering** by comparison. The core thesis of this chapter is that the most essential change in the educational-robot industry chain relative to 2019 is the emergence, from nothing, of **an independent "foundation-model / embodied-AI algorithm layer (VLA)"** — it crystallizes, as the robot's "brain," the end-to-end motion intelligence that "maps visual input directly to joint-control output," together with the educational large models and subject agents layered on top, and becomes the new value high ground ([212], [213]). At the same time, the upstream bottleneck has expanded from "chips" to "joints — harmonic drives — dexterous hands — six-axis force/tactile sensing," the content layer has upgraded from "course content" to "courses + practical training + certification," and operations and service have moved from an implicit cost to an explicit value layer.

The argument proceeds from the 2019 seven-layer baseline (§11.2) through the four reconstructions of 2026 — the new VLA algorithm layer (§11.3), the new upstream bottlenecks (§11.4), and the explicit emergence of the content and service value layers (§11.5) — to a naming proposal (§11.6) and the policy drivers as a strong external variable (§11.7), before a summary and a look ahead to Chapter 12.

11.2 The 2019 Seven-Layer Baseline: A Chain Centered on the Body Hardware

The foundation we build upon is the seven-layer industry chain established in the 2019 edition and corroborated by third-party accounts, restated here in full ([2] Figure 5-1; [46], a seven-layer-chain account):

1. **AI-chip manufacturers** — the most upstream link, new in the 2019 edition relative to 2016, providing the compute base; 2. **hardware manufacturers** — integration of the body structure, transmission, sensing, and other hardware; 3. **system-platform developers** — operating systems, middleware, and development frameworks; 4. **application-service providers** — applications and services for teaching scenarios; 5. **content suppliers** — courses, lesson plans, and teaching content; 6. **system integrators** — integrating hardware, platform, and content into deliverable solutions; 7. **channel distributors** — distribution networks facing schools and consumers.

The inner assumption of this structure is that **intelligence is endogenous to the hardware and platform,** so no separate layer is needed for "cognitive capability." It is precisely this assumption that foundation models and embodied intelligence broke after 2022.

11.3 The Largest Structural Change of 2026: A New "Foundation-Model / Embodied-AI Algorithm Layer (VLA)"

This is a layer that did not exist at all in either the 2016 or the 2019 edition. Its basis comes from two mutually corroborating industry observations. The end-to-end motion large model / vision–language–action (VLA) has become the core technical form of the robot's "brain," directly carrying the "visual input → joint-control output" mapping ([212]); and, in the embodied-intelligence industry panorama, the algorithm / large-model link has been stripped out of the hardware to form an industry layer that can be invested in, supplied, and iterated independently ([213]). The technical evolution (RT-2 → OpenVLA → π0 / π0.5 → Helix → GR00T and so on) has been tagged for maturity one by one in Chapter 4 and is not restated here; what this chapter cares about is its meaning for **industry structure.**

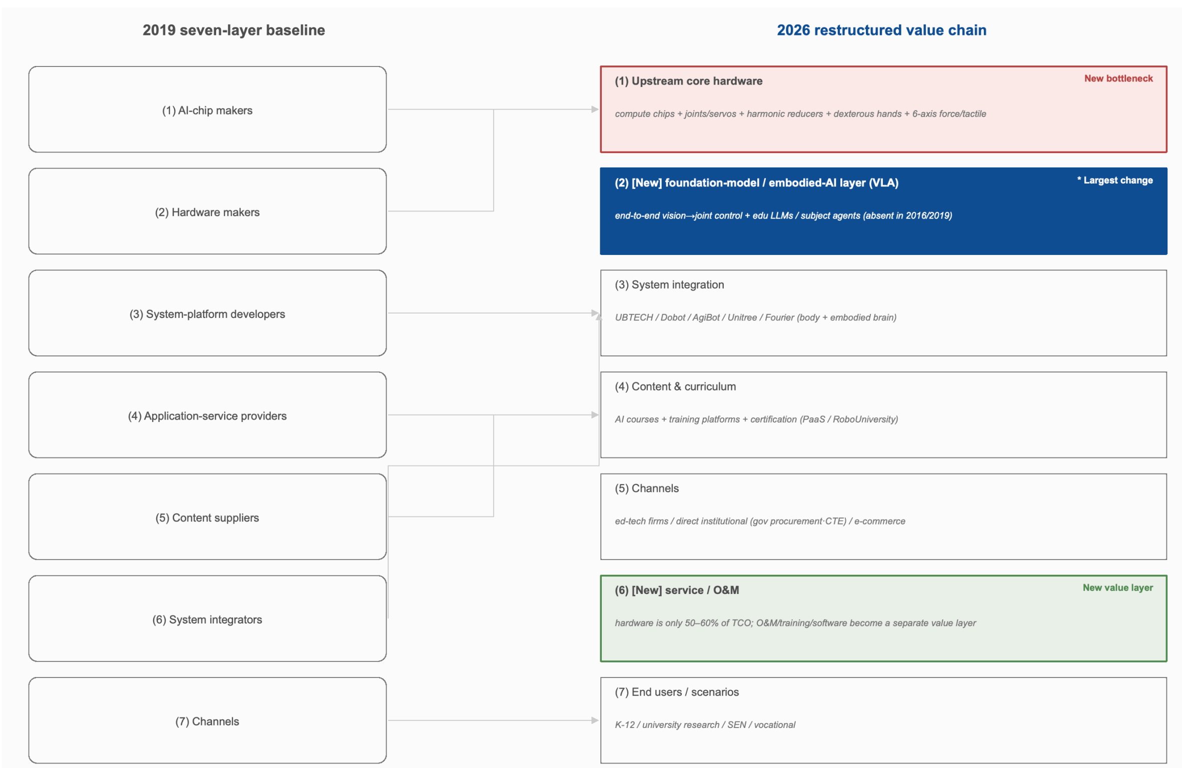
Why is this layer "from nothing" rather than "a refinement of an existing link"? Three reasons:

- **Priceability:** in the 2019 framework, "intelligence" was diffusely embedded in chips and platforms, with no independent deliverable or unit of charge; the VLA / educational large model can be priced separately through model licensing, API calls, agent subscriptions, and Model-as-a-Service (MaaS) — it has acquired the economic precondition to become an independent industry layer. - **Supplyability:**

foundation models and embodied algorithms are developed and supplied by dedicated algorithm vendors, and whole-robot makers can "buy the brain" rather than "build the brain" — precisely the mark of a new tier of industrial division of labor. - **Iterability:** the iteration pace of the model layer (monthly) is far faster than that of the hardware layer (yearly), and once decoupled the two evolve independently — the architectural implication of "brain ↔ body separation" (Chapter 5) manifests, at the industry level, as the stratification of the algorithm layer and the whole-robot layer.

In the educational context, this layer carries an additional overlay: atop the general-purpose VLA / foundation model, **educational large models and subject agents** (such as a Socratic tutor, a formative assessor, a learning-data analytics advisor) are further layered, so that the "brain" possesses both general motion/language intelligence and teaching-specific capabilities. This shares its source with Chapter 3's re-definition of "role orchestration" and with the framework argument for embodied intelligence empowering education ([47]).

A conservative flag is required: listing this layer separately is this blue book's **2026 reconstruction judgment** based on multi-source observation, not a conclusion from the original white papers; but its structural fact — "the algorithm / large model has been stripped from the hardware into an independent industry link" — has cross-supporting evidence from industry reports ([212], [213]).



Sources: SLIBNU 2019 seven-layer baseline (Fig. 5-1); 2026 restructuring from Qianzhan IER value-chain map + embodied-AI industry landscape + VLA industry report + humanoid-education O&M cost study (2026). Blue = new VLA layer; red = upstream new bottleneck; green = new value layer.

Figure 14.

11.4 Upstream Hard Bottlenecks: From "Chips" Forward to "Reducers — Screws — Motors — Force Sensing — Dexterous Hands"

The 2019 edition listed the AI chip as the most upstream bottleneck, reflecting the era's "compute scarcity." In 2026, as humanoids and high-degree-of-freedom bodies entered the industry's field of view, the upstream bottleneck **moved markedly forward and diversified:** beyond compute chips, **precision transmission (harmonic drives, planetary roller screws), joint drive (frameless torque motors), force/tactile sensing (six-axis force sensors, electronic skin), and the dexterous hand as the core of humanoid manipulation,** together with batteries and energy, jointly constitute the new upstream core hardware ([212], [213]). Humanoidization is the direct driver of this shift: as the body moves from "wheeled + fixed structure" to "high-degree-of-freedom humanoid," the precision, cost, and reliability of the transmission–drive–sensing chain become the true ceiling on whole-robot capability, and the localization progress and cost share of these links will, to a large degree, determine the affordability of the downstream whole robot — and thereby the underlying fulcrum of educational accessibility: "whether a university lab / vocational training room can afford an embodied platform."

This chapter only positions these bottleneck links structurally within the industry (the item-by-item detail of hardware parameters and cost curves belongs to Chapter 6); along the two dimensions of "cost share × localization progress," the upstream core components can be arranged into a **spectrum from the hardest bottleneck to the localization leader** ():

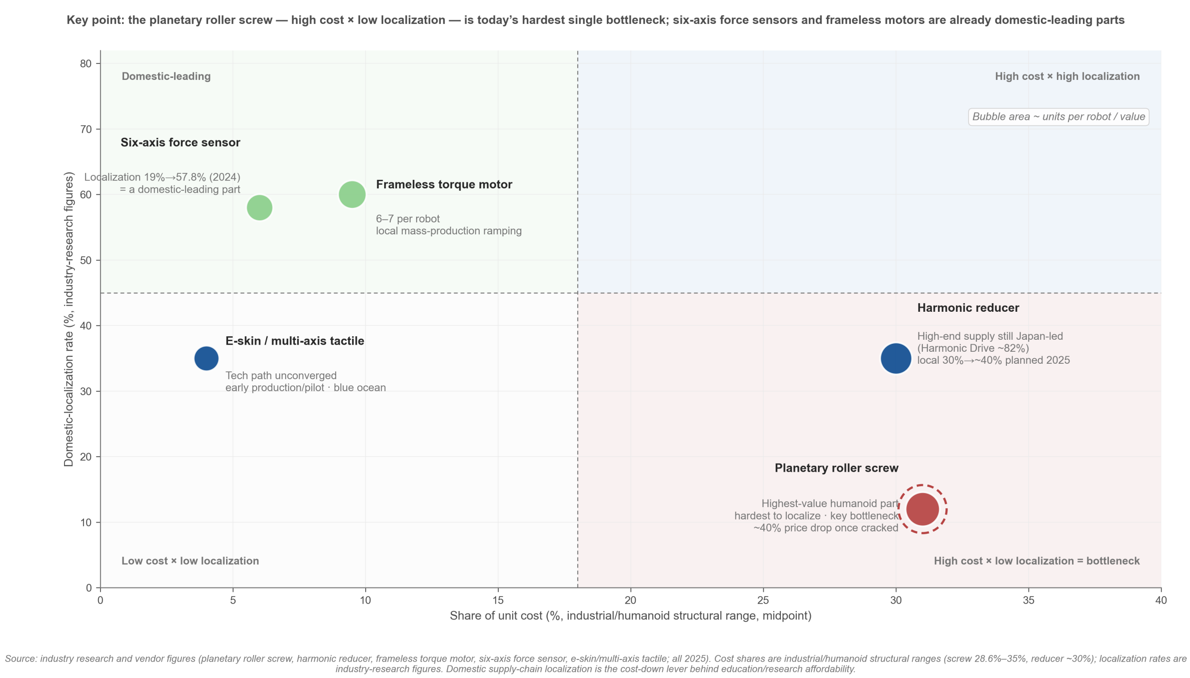


Figure 27.

Upstream core component	Cost share	Localization progress	Bottleneck position	Source
Planetary roller screw (core of the	highest, about 28.6%–35%	lowest, hardest to localize — domestic	the hardest bottleneck at present	[125]

linear actuator)		breakthrough under way, scale mass production not yet complete; price expected to drop about 40% after the breakthrough	— both the most expensive and the hardest to localize	
Harmonic drive (core of the rotary joint)	about 30% (the reducer is the top cost share among industrial-robot components)	Harmonic Drive Systems holds about 82% globally, Leaderdrive about 7%; localization rate about 30% → a planned 40% in 2025	the high end is still Japanese-led; localized mass production exists but the share is still small	214
Frameless torque motor (joint drive)	6–7 units needed per robot, about RMB 2,000–5,000 each	localized and already in mass production, ramp imminent (9 listed firms positioned)	localization advancing, with considerable room to bring prices down	215
Six-axis force sensor (core of force feedback)	unit price RMB 10,000–40,000 → toward the thousand-yuan level	localization rate rose from 19% to 57.8% (2024), with Bota Systems holding about an 80% share	a localization-leading link (already in mass production, ramping)	126
Electronic skin / multi-dimensional tactile	global market about US$7.1 billion → US$20.9 billion	early mass production / pilot (the technical route has not yet converged)	a blue-ocean early stage, route undecided	216

The core judgment of this spectrum is: **the decline of the whole-robot cost curve is bottlenecked not at the "localization-leading links" but at the "most expensive and hardest-to-localize" planetary roller screw and high-end harmonic drive.** The planetary roller screw has the highest value share among humanoid components (about 28.6%–35%), yet is the link with the lowest localization rate and incomplete scale mass production, and is therefore the **hardest single bottleneck at present** — it directly determines the cost of the linear actuator of a full-scale humanoid and thereby suppresses the pace of whole-robot cost reduction ([125]). The harmonic drive, though its localization rate has risen to about 40%, is still globally dominated by Harmonic Drive Systems (about 82%), with high-end supply constrained ([214]). In contrast are the two ends where localization already leads: the six-axis force sensor's localization rate leapt from 19% to 57.8% over four years ([126]), and the frameless torque motor is localized and in mass production with a ramp imminent ([215]).

The dexterous hand, as the core bottleneck of humanoid manipulation, is layered atop the transmission–sensing spectrum above. It is both the capability ceiling for high-degree-of-freedom manipulation and an integration of upstream components (roller screws + linkages + coreless motors + multi-dimensional tactile); Chinese vendors have already achieved a marked downgrade against high-end imports on tactile density and price (the detailed parameters belong to Chapter 6), but a gap remains between "high-degree-of-freedom demonstration" and "reliable mass production on the line." Its industry implication is that the cost and reliability of the dexterous hand are strongly coupled to the localization progress of its upstream screws, motors, and tactile sensing — **only when the upstream hard bottlenecks loosen will the cost curve of the dexterous hand and the whole robot move down in step.**

A conservative flag is required: the cost shares above are interval values on the industrial/humanoid structure basis (planetary roller screw 28.6%–35%, harmonic drive about 30%), and the share differs under different measurement bases; this chapter uses them only **qualitatively for structure** — "which link is most expensive, which is hardest to localize." The localization-rate figures follow industry-research bases and reflect substitution progress, not precise market share.

11.5 Value Redistribution: Content Upgraded to "Courses + Training + Certification," Operations Made Explicit as an Independent Value Layer

The third change of the 2026 reconstruction occurs in the **downstream value distribution** of the chain.

Content layer: from "course content" to "courses + a training platform + a certification system." The 2019 edition's "content suppliers" referred mainly to courses and lesson plans; in 2026 the content layer upgrades to an integrated "AI courses + training platform + certification system" — typical forms include UBTech's Xingzhi PaaS, the RoboUniversity-style platforms universities co-run with Unitree, and the course systems developed to accompany the general-education guidelines ([213]). Its industry implication is that content is no longer a "software giveaway" attached to hardware but an independent value link that carries certification credentials and can charge on a sustainable basis.

Service / operations layer: from an implicit cost to an independent value layer. A cost-structure fact worth stressing is that, in educational deployment, **hardware often accounts for only about 50%–60% of total cost of ownership (TCO),** with the rest made up of operations, training, software subscriptions, and other services — especially so for social robots ([151]). This means "selling a robot" is only the starting point of the value chain, and sustained operations and service are the body of long-term value. This chapter therefore makes "service / operations," implicit in the "application service / integration" position of the 2019 framework, **explicit as an independent value layer.**

A conservative flag is required: the 50%–60% ratio comes from a single media source ([151]), and this blue book uses it only as **qualitative evidence of cost structure** — "hardware is not all of TCO, and the operations share is considerable" — not as a precise financial ratio.

11.6 Naming Proposal: The Foundation-Model and Embodied-AI Service Layer

Synthesizing §11.3–§11.5, this report proposes formally naming the new algorithm layer **"the foundation-model and embodied-AI service layer"** — it encompasses foundation models, the VLA, agent frameworks, and MaaS, a structure entirely absent from both the 2016 and 2019 editions (the theoretical anchoring is in [47], sharing its source with Chapter 3's "adding a layer" re-definition).

This naming is a **[2026 extension]** proposal of this blue book, neither the original white papers' wording nor an industry-standard term; the purpose of proposing it is to provide, for subsequent statistical bases and policy formulations, an industry-layer concept able to accommodate new business forms such as "AI-education-agent subscriptions / MaaS / embodied-robot-as-a-service" — business forms that the 2019 edition's "four-model" market basis cannot accommodate (this basis limitation will be developed further in Chapter 12's look-back).

In one sentence, the evolution of the industry chain from 2019 to 2026: **from "a seven-layer chain centered on the body hardware" to "a chain with the embodied large model as the brain, joints / reducers / dexterous hands / force sensing as the new upstream bottlenecks, and course certification and operations service as new value layers."**

11.7 Policy Drivers: The "Action Plan — Base Schools — Local Targets" as a Strong External Variable

Unlike in 2019, the 2026 value-chain reconstruction occurs in an environment of **strong policy pull,** so policy becomes an external variable that cannot be ignored (the clause-level detail belongs to Chapters 2 and 14; this chapter only maps it to the industry):

- **National action plan:** the *"AI Plus Education" Action Plan* (Jiao Ke Xin [2026] No. 1) proposes accelerating the popularization of AI education in primary and secondary schools, building an educational intelligent-compute platform and grade-specific educational large models, and establishing a safety-review mechanism for educational large models ([19]); its higher-level policy, the *Opinions of the State Council on Deepening the Implementation of the "AI Plus" Initiative* (Guo Fa [2025] No. 11), sets targets of an agent penetration rate above 70% by 2027 and above 90% by 2030 ([17]) — directly creating scale demand for the "foundation-model and embodied-AI service layer." - **Base-school layout:** 509 AI-education base schools have been laid out nationwide, covering pilots across 7 eastern provinces, 20 central-and-western prefecture-level cities, and 18 universities ([20]) — a policy-driven expansion of the channel layer (direct institutional procurement). - **Local industry targets:** Shenzhen proposes more than 1,200 embodied-intelligence firms and an associated industry scale above RMB 100 billion by 2027, and Shanghai proposes an embodied core industry above RMB 50 billion by 2027 ([33], [213]) — injecting capital and capacity into the chain from both the upstream and the whole-robot-integration ends.

A conservative flag is required: the targets and base-school figures above follow the policy original; the local industry targets are planning figures that reflect policy intent rather than accomplished capacity, and should not be read as the measured scale of an educational-robot sub-market.

11.8 Chapter Summary

This chapter has given the 2026 reconstruction judgment of the educational-robot industry chain: relative to the 2019 seven-layer baseline, the largest structural change is the **addition of an independent "foundation-model / embodied-AI algorithm layer (VLA)"** — emerging from nothing to become the whole robot's "brain" and the new value high ground; the upstream bottleneck has moved forward from "chips" and expanded into "reducers — screws — motors — force sensing — dexterous hands," among which **the planetary roller screw (highest cost share, about 28.6%–35%, lowest localization rate) is the hardest single bottleneck at present,** the high-end supply of harmonic drives is still Japanese-led, while the six-axis force sensor (localization rate 19% → 57.8%) and the frameless torque motor are already localization-leading links (Fig. 27); the content layer has upgraded to "courses + training + certification," and service / operations has moved from an implicit cost to an explicit independent value layer; and this report proposes naming the new layer "the foundation-model and embodied-AI service layer" ([2026 extension]). Strong policy pull (the action plan / base schools / local targets) is the external variable of this reconstruction.

Two cautionary points bear restating. The new VLA layer is a structural judgment, and its downstream educational landing is still constrained by the "capability–deployment gap" revealed in Chapter 4 — an active algorithm layer does not equal synchronized maturity of classroom deployment. And figures such as the operations share and the local targets are, respectively, a single-source basis and a planning basis, used here only qualitatively.

The value-chain reconstruction sketches the skeleton of "how value is being redistributed," but the flesh on the skeleton — **how large the market is, how fast it grows, and whether historical forecasts have come true** — needs a dedicated chapter to compute. Chapter 12 turns to market size and TAM: characterizing the global and Chinese markets with conservative, multi-source ranges, and taking an honest look back at the 2019 edition's "US$84.1 billion" forecast — the latter confirming precisely §11.6's point that "the old basis can no longer accommodate the new business forms." Vendor-level financials, shipments, and the capital "ice and fire" are left to Chapter 13.

Chapter 12 Market Size and TAM: Conservative, Multi-Source Ranges

Scope and de-duplication note. *This chapter is the* ***master account*** *of the blue book's market figures — the size, growth, and definitional conflicts of the global and Chinese educational-robot markets, together with an honest look back at the 2016 and 2019 historical forecasts, are concentrated here. The sizing of the individual sub-tracks (STEM kits, social/companion, language, early-childhood, special education) is in Chapter 8; vendor financials and shipment rankings are in Chapter 13; humanoid-body hardware parameters are in Chapter 6. Where those topics are referenced, they are mapped as background, not recomputed. > >* ***Iron rules of scope (this chapter's highest priority):*** *(1) the global market is stated as a* ***range*** *(2026 about US$2.0–2.8 billion; around 2030 about US$5.5–5.8 billion; CAGR about 15%–29%), and* ***the most aggressive single CAGR of 28.8% is not adopted as the sole basis;*** *(2) the Chinese market's* ***broad-scope RMB 13.24 billion vs. narrow-scope RMB 4.4 billion are given together and explained, differing by nearly threefold, and must never be mixed;*** *(3) product form and the largest education level reach opposite conclusions across firms and must be flagged as conflicts; (4) the humanoid / quadruped / dexterous-hand totals are all* ***not education TAM*** *and serve only as background for educational positioning — the humanoid 2026 shipment forecast, raised by Morgan Stanley to about 50,000 units, must be flagged "an institutional forecast with large basis swings," the dexterous-hand US$815 million vs. US$1.48 billion dual basis presented side by side with the 40% CAGR treated as an aggressive single source, and the US$5-trillion-by-2050 figure used only as a footnote; (5) the historical "US$84.1 billion" is flagged "a forecast, already an over-estimate, with an over-broad basis" and* ***must never be reused as accomplished fact.***

12.1 Introduction: Why Market Figures Must Be Presented Side by Side Rather Than Reduced to One Value

The market size of educational robots is the class of figure most easily misused in this blue book, for three converging reasons. Different firms' **definitional bases for "educational robot" differ enormously** — whether STEM/coding kits are included, whether humanoid teaching robots are included, whether training services and consumer toys are included can make the same year's market size differ severalfold. Most reports are published as **forecasts,** besides, and the divergence in forecast CAGR (from 14.67% to 28.8%) magnifies over a few years into a vast divergence in scale. And with Chinese- and English-language sources and broad and narrow scopes coexisting, it is easy to mix and splice them inadvertently, manufacturing a figure that "looks precise but is in fact distorted."

This chapter's methodological stance is therefore clear: **present conflicting bases side by side, explain the differences, and take a conservative range rather than the most aggressive single value.** This stance is itself the methodological highlight of the 2026 edition relative to its predecessors — and its cautionary tale is precisely the 2019 edition's "US$84.1 billion" forecast (§12.5).

The chapter proceeds through the global market (§12.2), the two Chinese scopes (§12.3), the conflicts of region and sub-form (§12.4, including the humanoid-total background), and the look back at historical forecasts (§12.5).

12.2 The Global Educational-Robot Market (Narrow Body Scope, a Multi-Source Range)

The bases of the main firms, placed side by side ():

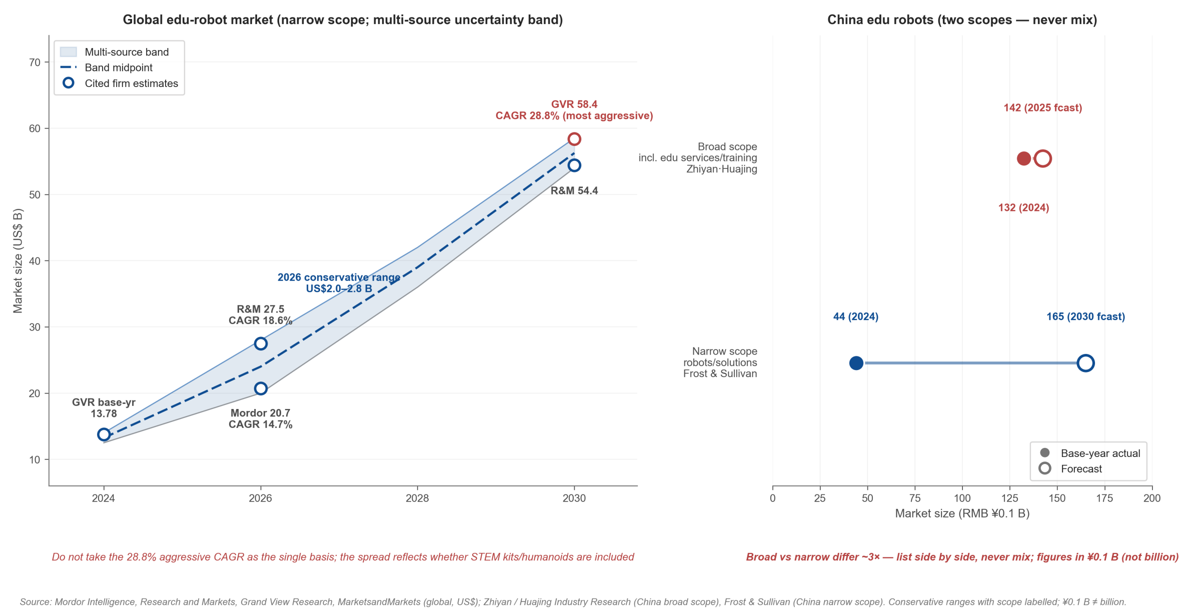


Figure 15.

Firm	2026 size	2030 / terminal size	CAGR	Scope notes
Mordor Intelligence	about US$2.07 billion	2031 about US$4.11 billion	14.67%	from a 2025 base of US$1.81 billion, published 2026-06 ([160])
Research and Markets	about US$2.75 billion	2030 about US$5.44 billion	18.6%	structure close to GVR ([16])
Grand View Research	— (2024 base US$1.378 billion)	2030 about US$5.84 billion	28.8% (most aggressive)	report page 403; cross-checked three ways via press release + PR Newswire ([15], [217], [218])
MarketsandMarkets	— (2022 US$1.4	—	17.3%	includes Service +

	billion → 2027 US$3.2 billion)			Industrial ([219], [220])
DIResearch (Baijian Fanglue)	— (RMB basis)	2032 about RMB 36.83 billion	13.80%	global RMB basis, single firm ([221])
MarketResearch.Biz (Chinese re-citation)	— (2024 about US$2.4 billion)	2030 about US$5.8 billion	26.44%	a single Chinese re-citation, treated with caution ([222])

The conservative conclusion to adopt (global): synthesizing the table above, this blue book uniformly uses a **range** in the body text —

> *The global educational-robot market (narrow body scope) is about **US$2.0–2.8 billion** in 2026 and about **US$5.5–5.8 billion** around 2030, with a CAGR of about **15%–29%.***

The divergence between the lower bound (Mordor US$2.07 billion / 14.67%) and the upper bound (GVR US$5.84 billion / 28.8%) arises **mainly from whether STEM/coding kits and humanoid teaching robots are included,** as well as from differences in base year and forecasting method. Unpacking the divergence firm by firm shows why "side by side rather than a single value" is the only cautious course. Mordor counts from a 2025 base of US$1.81 billion at a 14.67% CAGR, giving a comparatively conservative path of about US$4.11 billion by 2031 ([160]); Research and Markets gives about US$5.44 billion by 2030 at an 18.6% growth rate, with a structure close to GVR's ([16]); MarketsandMarkets, because it includes Service and Industrial scopes, gives a broader path of US$1.4 billion in 2022 → US$3.2 billion in 2027 at a 17.3% CAGR ([219]); and DIResearch (Baijian Fanglue) forecasts about RMB 36.83 billion by 2032 on an RMB basis (single firm, [221]). Under the same "global educational robot" label, base year, growth rate, currency, and coverage differ pairwise, and any practice of adding them directly or selecting one as settled would manufacture a figure that "looks precise but is in fact distorted."

This blue book does not adopt the most aggressive CAGR of 28.8% as the sole basis, nor does it treat any single firm's point forecast as accomplished fact; GVR's US$5.84 billion serves only as the upper bound of the range, and since its report body page once returned a 403 error, the related figures are cross-confirmed by its press release and a third-party newswire ([217], [218]). MarketResearch.Biz's US$5.8 billion, being a single Chinese re-citation, serves only as corroboration ([222]). One sense of magnitude must be stressed: even at the upper bound, the global narrow-scope educational-robot-body market is under US$6 billion by 2030 — which, against the humanoid total (§12.4, on the order of tens of billions of dollars by 2030), confirms the whole paper's judgment that "education is a secondary scenario of embodied intelligence"; misreading the educational-robot market as being of the same magnitude as the humanoid total is another common conflation of scope.

12.3 The Chinese Market: Two Scopes Side by Side, Never to Be Mixed

The Chinese educational-robot market has two scopes that **differ by nearly threefold,** and the two **must be given together and explained, never mixed or spliced:**

Scope	2024 size	Subsequent	Firm	Coverage
Broad scope	about RMB 13.24 billion	2025 projected about RMB 14.24 billion	Zhiyan Consulting, Huajing Industry Research	includes robotics-education services / training ([223], [224])
Narrow scope	about RMB 4.4 billion	2030 about RMB 16.5 billion	Frost & Sullivan (cited by Qianzhan)	body / solutions ([159])

The difference between the two scopes is **not a data error but a difference in definition:** the broad scope folds "robotics-education services, training-institution revenue" and the like into the market basket, making it markedly larger (on the Zhiyan basis, the Chinese market was RMB 380 million in 2015, RMB 9.26 billion in 2020, and about RMB 13.24 billion in 2024, [223]); the narrow scope counts only educational-robot bodies and solutions, at about a third of the broad scope.

The conservative adoption (China): the body text **presents the two scopes together and explicitly explains the difference.** If a comparable conservative anchor is needed, the narrow scope's **about RMB 16.5 billion in 2030** is recommended (Frost & Sullivan, [159]), because its scope (body / solutions) is more comparable to the global narrow-body scope. Any practice of directly comparing the growth of "broad-scope RMB 13.24 billion" with "narrow-scope RMB 4.4 billion," or computing it side by side with the global body scope, is a mixing of scopes, which this blue book uniformly avoids.

12.4 Region, Product Form, and Education Level: Opposing Conclusions Must Be Flagged

Even within the same "global body" basket, different firms' judgments of **structural shares** are in direct conflict and must be flagged side by side (no single one chosen as settled):

- **Regional distribution:** GVR says North America held 35.6% in 2024 ([15]); Mordor says Asia-Pacific held 34.40% and was the largest region, North America 28%, Europe 24% ([160]). The **safely adoptable qualitative conclusion** is "an Asia-Pacific-and-North-America dual core, with Asia-Pacific growing faster"; which is the largest region differs by basis, so the two must be presented side by side. - **Product form (opposing conclusions):** GVR says **non-humanoid-led** (non-humanoid about US$932 million, [15]); Mordor says **humanoid holds 49.13%** ([160]). This is a pair of **opposing** judgments, arising from a basis divergence over whether "educational robot" is primarily coding kits or humanoid teaching machines, and **the conflict must be flagged, with neither chosen. - Largest education level (opposing conclusions):** GVR says secondary holds 39.3% ([15]); Mordor says primary holds 36.81% ([160]). Likewise handled side by side.

The humanoid-robot total (not education TAM, background only). To give a sense of scale for "education is a secondary scenario of embodied intelligence," the humanoid total is listed here, but a strict declaration is required: **this is not educational-robot TAM; education is only one of the downstream scenarios of the humanoid.**

- Shipments (2025 measured): global humanoid shipments about **13,000 units** (Omdia, [225], [226]) vs. about **18,000 units** (IDC, about +508% year over year, [227]) — **the range "13,000–18,000 units" is adopted side by side,** with Chinese manufacturing about 87% (the shipment-vendor breakdown belongs to Chapter 13). - Shipments (2026 forecast, large basis swings, presented side by side): in June 2026 Morgan Stanley **raised** its forecast of that year's global humanoid shipments **to about 50,000 units** (double the prior value of about 28,000; the firm's January forecast had been only about 14,000, [228]) — within half a year the same firm's basis jumped successively from 14,000 → 28,000 → 50,000 units, a textbook case of "an institutional forecast with large basis swings," so **the body text gives only a trend indication and adopts no single point value as settled;** the same firm also raised its China 2030 shipment forecast to about 446,000 units, with the China humanoid market about US$2 billion in 2026 → about US$15 billion in 2030 ([228]). - Market size: the global humanoid market about US$3 billion in 2025 → about US$28 billion in 2030 (Morgan Stanley); Goldman Sachs estimates about US$38 billion in 2035 ([229], [230]) — single-firm forecasts, basis flagged. - Ultra-long-range: Morgan Stanley says the global humanoid market may reach **US$5 trillion** by 2050 ([231]) — **this figure is used only as a footnote, not asserted in the body text.**

The quadruped-robot total (not education TAM, background only). The quadruped is the most realistic platform for embodied intelligence to enter education at present (vocational/university training, see Chapter 9), but its market size is likewise **not education TAM:** in 2025 the global quadruped-robot market was about **US$541 million with sales of about 18,500 units** (an average price of about US$32,000 per unit) → about US$1.778 billion in 2034, a CAGR of about 18.7% ([141], a single-firm basis); within it, Unitree's cumulative quadruped sales exceeded 30,000 units, first in the world (prospectus basis, [107]). Two points of scope-background guard against extrapolation. Chinese robot-dog sales are growing rapidly (about 23,300 units in the first three quarters of 2024, about +72.22% year over year, with one firm further forecasting about 394,800 units in 2031), but that 394,800-unit figure is a **single firm's long-range forecast,** which must be flagged "large basis divergence" and not taken as accomplished fact. And the reason the quadruped is "one of the few closure points of the capability–deployment gap" lies in its **low price threshold for entering education** — the Unitree Go2 Air at about US$1,600, the Go2 EDU standard edition at about US$5,990, and the DEEP Robotics Jueying Lite3 education/research edition at about US$2,890 are an order of magnitude below the humanoid G1 EDU's about US$43,900–73,900, which is the economic root of why the quadruped has actually entered Chinese vocational/university classrooms while the humanoid still mostly stays on research platforms (see the HZPT case in Chapter 9). But an accessible price does not mean a large market: the quadruped total of US$541 million is still far smaller than the narrow-scope educational-robot-body market itself, and must certainly not be conflated with the humanoid total.

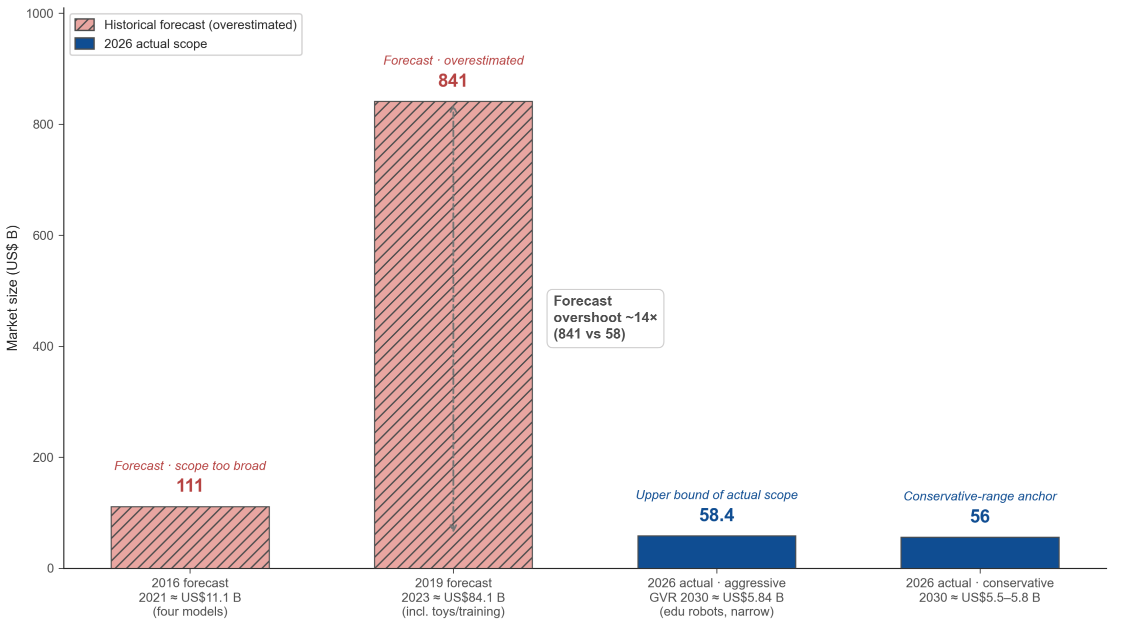
The dexterous-hand market (a new upstream humanoid bottleneck, not education TAM). As the core bottleneck of humanoid manipulation (the structural position is in Chapter 11), the dexterous-hand market has enormous basis divergence and must be presented side by side and treated with caution: one basis for the global multi-finger dexterous-hand market is about **US$815 million** in 2024 → about US$10.3 billion in 2031, a CAGR of about **40.4%** (Valuates, [232]); another basis is about **US$1.48 billion** in 2025 (360iResearch, [233]). **The two bases differ by nearly a factor of two, the 40% CAGR is a single firm's aggressive basis, and the body text uses this only as "background for the upstream cost / capability curve," not as education TAM, nor does it adopt the aggressive CAGR as settled.** The dexterous hand's relation to education is **indirect and upstream:** it determines the ceiling and cost curve of humanoid manipulation, and the rapid breakthroughs of Chinese vendors in tactile density and price (such as the Linker Hand L20 from about RMB 6,666 and the OYMotion ROHand from about RMB 8,888, an about 1/20-magnitude down-step from the Shadow Hand's about US$74,000) are lowering the embodied body's capability-acquisition cost from upstream, indirectly affecting the final price at which "a university lab can afford" one; but the dexterous hand is itself a research / data-collection platform, **not a primary-or-secondary classroom teaching aid,** nor an education-market size in any sense.

The "basis swings" of the humanoid 2026 shipment forecast warrant special warning. The same firm's (Morgan Stanley) forecast of 2026 global humanoid shipments jumped, within half a year, from about 14,000 units in January to 28,000 units and then up to about 50,000 units in June ([228]) — successive doublings that are themselves the most direct evidence of "extremely high institutional-forecast uncertainty." If such a figure is excerpted at one point in time and cited as "fact," it can easily mislead; this blue book uniformly gives only a trend indication and adopts no single point as settled. Market size is the same: the global humanoid market about US$3 billion in 2025 → about US$28 billion in 2030 (Morgan Stanley) and about US$38 billion in 2035 (Goldman Sachs, [229]) are all single-firm forecasts; as for the ultra-long-range US$5-trillion-by-2050 figure, it is used only as a footnote and never enters the body text as an assertion.

A **layered, conservative TAM picture** can thus be sketched: the innermost layer is this chapter's object of computation — the global narrow-scope educational-robot-body market (about US$2.0–2.8 billion in 2026; about US$5.5–5.8 billion around 2030); outside it, the broad/narrow dual scope of the Chinese market (RMB 13.24 billion / RMB 4.4 billion, to be presented side by side); and outside that, the humanoid/quadruped/dexterous-hand totals as positioning background (all not education TAM, institutional forecasts, with large basis divergence). The three layers **must not be added or spliced across layers,** or one falls into scope-mixing. To restate: the 2026/2030 forecasts of the humanoid, quadruped, and dexterous-hand totals are all flagged "an institutional forecast with large basis divergence (the humanoid 2026 shipment basis swings especially sharply)"; what they provide for educational robots is **positioning background** (the humanoid cost coming down, the quadruped platform becoming accessible, the dexterous-hand upstream curve, the supply-chain configuration), not the education-market size itself.

12.5 An Honest Look Back at Historical Forecasts: From "US$84.1 Billion" to "≤ US$5.8 Billion"

This blue book folds the look back at its predecessors' market forecasts into the body text as a **methodological self-check** — one of the key improvements of the 2026 edition over the 2016 and 2019 editions ().



Source: SLIBNU Educational Robot White Paper 2016 Table 5-1 (US11.1B), 2019Table6 – 1(US84.1 B, four models incl. consumer toys/training); Grand View Research 2030 (US5.84B, narrowscope). TheUS84.1 B figure was a research-team forecast with overly broad scope and is clearly overestimated — it must not be reused as established fact.

Figure 16.

The two historical forecasts:

- **The 2016 edition** forecast: by 2021, the global educational-robot market at about **US$11.1 billion** (the sum of four models — terminal consumer / educational institution / educational kits / STEAM toys, [1]; corroborated by ScienceNet's "may reach tens of billions of dollars in five years" report, [234]). - **The 2019 edition** forecast: by 2023, the global market at about **US$84.1 billion** (the same four models, including a large volume of consumer toys and training services, [2]; corroborated by Sina's "six core viewpoints" report, [235]).

Against the 2026 measured figure: even using the **most aggressive** GVR basis of §12.2, the global educational-robot market reaches only about **US$5.84 billion** by 2030 ([15]). In other words, the 2019 edition's "US$84.1 billion by 2023" forecast is more than an order of magnitude higher than the 2030 scale under the most optimistic 2026 basis.

Basis diagnosis and the conclusion to adopt: US$84.1 billion is **a project team's forecast value, with an over-broad basis** — it folds a large volume of consumer toys and training services into the "educational robot" basket and therefore **markedly over-estimates** the narrow-scope educational-robot-body market. This deviation yields three drafting disciplines:

1. **Explicit flagging:** "US$84.1 billion / US$11.1 billion" is uniformly flagged "a forecast value, already an over-estimate, with an over-broad basis," and **must never be reused as accomplished fact;** 2. **Basis separation:** the "narrow-scope educational-robot-body market" (the object of computation in §12.2/§12.3) must be distinguished from the "general robotics-education / STEAM consumer market" (a broader basis, easily mixing in training and toys); 3. **A methodological highlight:** this basis lesson is turned into the 2026 edition's statistical discipline of "multi-source side by side, conservative range, basis separation" — which echoes Chapter 11 §11.6's point that "the old four-model basis can no longer accommodate new business forms such as AI-education-agent subscriptions / MaaS / embodied-robot-as-a-service": part of the root of the historical over-estimate is precisely that the old basis both mixed in consumer goods that should not be counted and could not characterize the newly emerging service-ization business forms.

12.6 Chapter Summary

This chapter has given the master account of the educational-robot market and held to the iron rules of scope: the global narrow-body market is about **US$2.0–2.8 billion** in 2026 and about **US$5.5–5.8 billion** around 2030, with a CAGR of about **15%–29%** (a range, not the most aggressive 28.8%); the Chinese market's **broad scope of about RMB 13.24 billion and narrow scope of about RMB 4.4 billion are given together and explained, differing by nearly threefold and never to be mixed,** with the narrow scope's about RMB 16.5 billion in 2030 serving as a conservative anchor; region, product form, and largest education level reach opposing conclusions across firms and have been flagged as conflicts side by side; the humanoid / quadruped / dexterous-hand totals are all not education TAM and serve only as background — the humanoid 2026 shipments raised by Morgan Stanley to about **50,000 units** (flagged "an institutional forecast with large basis swings"), the quadruped about 18,500 units / US$541 million in 2025, the dexterous-hand US$815 million vs. US$1.48 billion presented side by side (the 40% CAGR an aggressive single source treated with caution), and the US$5 trillion / 2050 used only as a footnote; the 2019 edition's "US$84.1 billion" is flagged as an over-estimated forecast value and given an honest look back.

A cautionary restatement: all sizes in this chapter are **ranges or side-by-side bases,** and any practice of compressing them into a single precise value, or splicing across bases, departs from this blue book's statistical discipline. The market total answers "how big the cake is," but "who is slicing the cake, what each vendor's financials and shipments are, and why the capital is ice and fire" is a question of the competitive landscape. Chapter 13 turns to the competitive landscape: it handles, in concentrated form, the FY2025 financials of representative vendors (**holding strictly to the currency corrections: UBTech's 2025 total revenue of RMB 2.001 billion, AI education of RMB 413 million, Unitree's 2025 revenue of RMB 1.708 billion, iFLYTEK's smart-education revenue of RMB 8.967 billion — all in RMB 亿元 (100 million yuan), not billions of dollars**), the 2025 humanoid shipment rankings, the capability × vendor matrix, and the "ice and fire" contrast of "the financing chill for traditional educational robots and the bankruptcy of an established vendor vs. the capital frenzy for embodiment / humanoids."

Chapter 13 Competitive Landscape: Representative Vendors, Shipment Rankings, and Capital Divergence

Scope and de-duplication note. *This chapter is the* ***master account*** *of the blue book's vendor financials, shipment rankings, and the capital side. The vendors' education-deployment status and case detail belong to Chapter 9 (the deployment narrative is not repeated here, only the financials and shipments); humanoid-body hardware parameters belong to Chapter 6; the market-total TAM belongs to Chapter 12 (this chapter discusses single-firm financials and does not recompute the total); the product-parameter matrix is in Appendix E. > >* ***Iron rule of currency basis (this chapter's highest priority):*** *all Chinese company financials are in RMB 亿元 (100 million yuan), never in "billions." UBTech's FY2025 total revenue was RMB 2.001 billion (about US$280 million), AI education RMB 413 million, humanoid RMB 821 million; Unitree's 2025 revenue was RMB 1.708 billion; iFLYTEK's smart-education 2025 revenue was RMB 8.967 billion — all in RMB 亿元, none in billions of dollars.* *Overseas search engines repeatedly misread "X 亿元" as "X billion dollars" (reading RMB 413 million as 4.13 billion, RMB 1.708 billion as 17 billion); this chapter re-checks every financial figure against the source. RMB 100 million is about US$14 million (about US$0.014 billion), decidedly not 1 billion. > >* ***Other conservative caveats:*** *UBTech's education business must use the FY2025 audited basis of* ***RMB 413 million (up 13.7%),*** *and* ***must not use the old 2024 value of RMB 363 million*** *returned by early searches; shipment data differ between the Omdia basis and the prospectus/annual-report basis and must be presented side by side (Unitree about 4,200 units / second in the world on Omdia vs. >5,500 units / first in the world in the prospectus); Fourier's financing amount is a media basis, flagged PENDING; US$5 trillion / 2050 is not asserted in the body text; and it must always be stressed that "education is a downstream application of embodied intelligence, not its main arena."*

13.1 Introduction: Reading Vendors on a "Non-Education Main Arena"

Characterizing the competitive landscape of educational robots requires one premise to be stated up front: **education is not the main arena of today's embodied-intelligence / humanoid-robot vendors.** The leading vendors (UBTech, Unitree, Agibot) draw their revenue mainly from research, industry, commercial, and quadruped scenarios, and education (especially K-12) is only one of their downstream applications; the largest AI-education revenue in China by scale, in fact, comes from the **screen-based, software-hardware-integrated** smart education of iFLYTEK (2025 revenue of RMB 8.967 billion), an order of magnitude beyond the education segment of any embodiment vendor. When reading vendor financials, this chapter therefore gives both "total revenue" and "education-related revenue," and distinguishes "embodiment vendors" from "screen-based AI-education vendors," to avoid misreading a vendor's overall size as the size of the education market — consistent with Chapter 12's judgment that "education is a secondary scenario of the humanoid."

This chapter's financials uniformly use the **latest FY2025 audited annual reports / prospectuses** (replacing the 2024 old values in the source base). It first re-checks the representative vendors' financials (§13.2, where the currency correction is of the first importance), then presents the 2025 humanoid shipment rankings (§13.3, the Omdia basis and the annual-report/prospectus basis side by side), then gives the capability × vendor matrix (§13.4), the capital "ice and fire" contrast (§13.5), and the sustainability risk of incumbent vendors (§13.6, Aldebaran's bankruptcy), and finally handles large orders and valuations (§13.7) before a summary and look ahead.

13.2 Representative-Vendor Financials

The table below is the chapter's core; every figure has had its currency unit re-checked against the source, **"亿元" not "billion":**

Company	Key financials (RMB 亿元, re-checked)	Education-related	Source
UBTech (HKEX: 09880)	FY2025 total revenue RMB 2.001 billion (up 53.3% YoY); gross margin 37.7%; net loss RMB 790 million	AI-education business RMB 413 million (up 13.7%) (replacing the old 2024 value of RMB 363 million)	[166], [168], [163]
UBTech (humanoid segment)	2025 humanoid revenue RMB 821 million (up +2,203.7% YoY, 41.1% of total revenue), cumulative shipments 1,079 units, humanoid gross margin 54.6%	includes the research-and-education Tien Kung Walker	[166], [236]
Unitree	2025 revenue RMB 1.708 billion (up 335.36%), non-GAAP net profit RMB 600 million (up 674.29%), main-business gross margin 60.27%; 2024 revenue RMB 392 million; valuation about RMB 42 billion; STAR Market IPO planned raise RMB 4.202	roughly 3/4 of humanoids (74% in the prospectus) sold to universities for research, only about 9% entering actual industrial use	[107], [169], [170]

	billion		
iFLYTEK (002230)	2025 smart-education revenue RMB 8.967 billion (up 24.04%), about 33% of total revenue, the company's largest business segment; consumer AI learning machines up +104% in H1 2025	smart education = China's largest AI-education revenue base (screen-based, software-hardware integrated, not a humanoid body)	[237], [238]
Agibot (Zhiyuan)	achieved commercial mass production of a general-purpose robot in December 2024 (first in the industry); 2025 shipments 5,168 units	education is the entry-level "a hundred cities, ten thousand schools" plan (planning stage)	[213], [226]
Fourier	completed multiple financing rounds in 2024, with Sequoia China and others participating, total projected by media at >RMB 1 billion (PENDING, financing amount treated with caution)	entered via rehabilitation + a research-and-development platform; education is a developer / university platform	[143]

Item-by-item correction notes:

- UBTech FY2025 total revenue RMB 2.001 billion ≠ 20 billion. UBTech listed on the Hong Kong Stock Exchange on December 29, 2023, and is called "the first humanoid-robot stock"; its **FY2025 full-year total revenue was RMB 2.001 billion (about US$280 million), up 53.3% YoY,** with a gross margin of 37.7% and a net loss of RMB 790 million ([166], [168], re-checked against the source). Of this, **the AI-education business was RMB 413 million, up 13.7% YoY** ([163], [166]) — a special correction is needed here: **the audited-annual-report basis of RMB 413 million should be used, not the old 2024 value of RMB 363 million (up 4.6%, an old-year basis) returned by early searches;** RMB 413 million ≈ US$57 million, **decidedly not 4.13 billion.** This ratio (education at about 20.6% of total revenue) shows that even for a vendor as deeply invested in education as UBTech, education is only about one-fifth of total revenue, and the full-scale humanoid (RMB 821 million, 41.1%) has already overtaken education as the larger single segment. **- UBTech 2025 humanoid RMB 821 million ≠ US$8.2 billion.** Overseas it was once misread as US$8.2 billion; the correct value is RMB 821 million, 41.1% of total revenue, up +2,203.7% YoY, with cumulative shipments of 1,079 units and a humanoid gross margin of 54.6% ([166], [236]). **- Unitree RMB 1.708 billion ≠ 17 billion.** Unitree is in the midst of a STAR Market IPO (planned raise RMB 4.202 billion); **2025 revenue was about RMB 1.708 billion, up 335.36% YoY,** with non-GAAP net profit of RMB 600 million (up 674.29%) and a main-business gross margin of 60.27% (2024 revenue RMB 392 million; [107], [169], [170]). Overseas searches once misread it as "17 billion," which this chapter rejects. The prospectus makes clear, tellingly, that **roughly 3/4 (74%) of humanoids are sold to universities for research and only about 9% enter actual industrial use** (presented side by side with the source base's "73.6% from research and education" [the January–September 2025 window], rendered uniformly as "roughly three-quarters from research and education," [107]) — the most powerful financial corroboration of the judgment (Chapter 9) that "the humanoids that can enter education are mainly university/research platforms, not K-12 classrooms." - **iFLYTEK smart education RMB 8.967 billion ≠ 89.67 billion (new).** iFLYTEK's 2025 smart-education revenue was **RMB 8.967 billion, up 24.04% YoY,** about 33% of the company's total revenue and its largest business segment ([237], [238]); RMB 8.967 billion ≈ US$1.25 billion, **still "亿元" not billion.** Note its basis: this is **screen-based, software-hardware-integrated AI-education revenue** (learning machines, smart-classroom solutions), not a humanoid / body robot — it is China's largest AI-education revenue base, an order of magnitude beyond the education segment of any embodiment vendor (UBTech's AI education at RMB 413 million is only about 1/22 of it), which precisely illustrates the blue book's spine: "what enters classrooms at scale is still the screen-based software agent, not the robot embodiment." - **Fourier's financing amount flagged PENDING.** Fourier's >RMB 1 billion financing is a single, media-aggregated basis, and the amount needs secondary verification; this chapter flags it **PENDING-secondary** and does not cite it as accomplished fact ([143]).

13.3 2025 Humanoid Shipment Rankings (Omdia Basis, Side by Side with Annual Report / Prospectus)

The 2025 global humanoid-robot shipment rankings have **two opposing bases** — the Omdia third-party basis and the vendors' prospectus/annual-report basis — which must be presented side by side:

Vendor	Omdia basis	Vendor basis (annual report / prospectus)	Source
Agibot	5,168 units (first, about 39%)	same as Omdia	[226], [225]
Unitree	about 4,200 units (second in the world)	prospectus >5,500 units, first in the world, 32.4% share (G1)	[226], [107]
UBTech	about 1,000 units	annual report 1,079 units	[226], [236]
U.S. Tesla / Figure / Agility	about 150 units each	—	[226]

The basis difference must be presented side by side, not forced into convergence: Unitree's shipments carry a basis conflict in which **even the ranking is rewritten** — about 4,200 units and second in the world on the Omdia basis, vs. >5,500 units and first in the world (32.4% share) in the prospectus; UBTech likewise has 1,000 units on Omdia vs. 1,079 in the annual report. The two are **presented side by side;** if a unified basis is needed for cross-comparison, the Omdia third-party basis is recommended as primary, with the vendors' prospectus basis as corroboration, and "who is first in shipments differs by basis" explicitly flagged. On the structural facts: **Chinese manufacturing accounts for about 87% of global humanoid shipments, and the top three vendors together for about 78%** ([225], [226]) — China has formed a dominant position in whole-humanoid shipments. This remains, to restate the point, the shipment landscape of the **humanoid total:** about three-quarters of Unitree's >5,500 humanoids flow to university research platforms, with education only a very small downstream slice within.

13.4 The Capability × Vendor Matrix: A Comparison of Education-Entry Degree

Comparing the main vendors along the two axes of "core capability dimension" and "education-entry degree" yields a capability × vendor matrix (). This section gives only a qualitative comparison; deployment detail is in Chapter 9 and product parameters in Appendix E:

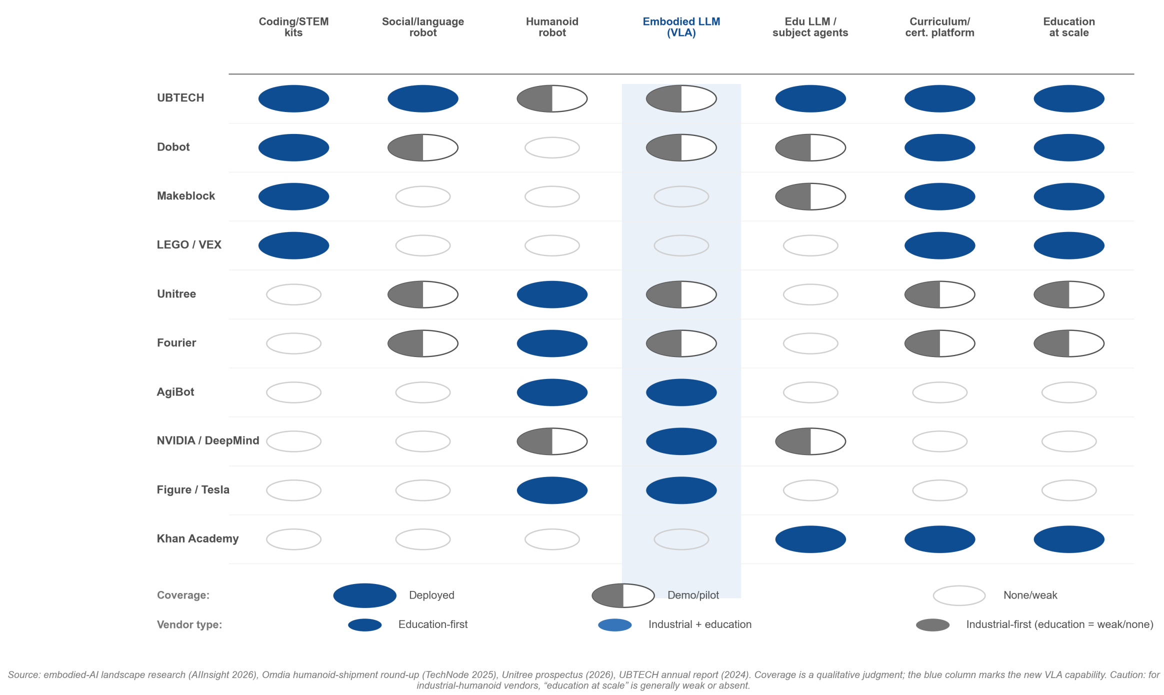


Figure 17.

- **UBTech:** the "education-native" vendor with the broadest capability coverage — from small coding robots (Alpha Mini / Wukong) to the research-and-education humanoid (Tien Kung Walker), with the Xingzhi PaaS providing courses and certification; **the mainstay of primary/secondary scaling is still the small coding robot** (FY2025 AI education RMB 413 million, covering nearly 2,000 public primary and secondary schools in China, a vendor basis), with the full-scale humanoid in classrooms still early ([163], [213]). - **Unitree:** enters the **university/research teaching platform** with a high-cost-performance humanoid (G1 EDU), the representative of the "affordable humanoid"; its education-entry degree concentrates in universities and research (roughly 3/4 of humanoids sold to universities for research), not the K-12 classroom ([107]). - **Agibot:** first in shipment volume, but with industry / commercial as the main body; education is the entry-level "a hundred cities, ten thousand schools" **plan,** at the planning/early stage ([213], [226]). - **Fourier:** enters via rehabilitation and a research-and-development platform; its education-entry degree concentrates in the developer / university platform ([143]).

The overall judgment from the matrix: on the capability dimension, the leading vendors' humanoid capabilities are converging fast; but on the **education-entry-degree** dimension there is a clear stratification — UBTech scaling on the K-12 coding track, Unitree and Fourier landing on the university/research platform, and Agibot still at the planning stage on the education side; **the full-scale humanoid entering K-12 classrooms at scale remains, to date, a concept / early stage.** This is entirely consistent with Chapter 6's hardware judgment that "primary/secondary scaling is still mainly small coding/social bodies" and Chapter 9's deployment judgment that "the humanoids that can enter education are mainly university research platforms."

13.5 Capital Divergence: The Chill for Traditional Educational Robots vs. the Frenzy for Embodiment / Humanoids

The 2024–2026 capital side presents a contrast of striking narrative tension ():

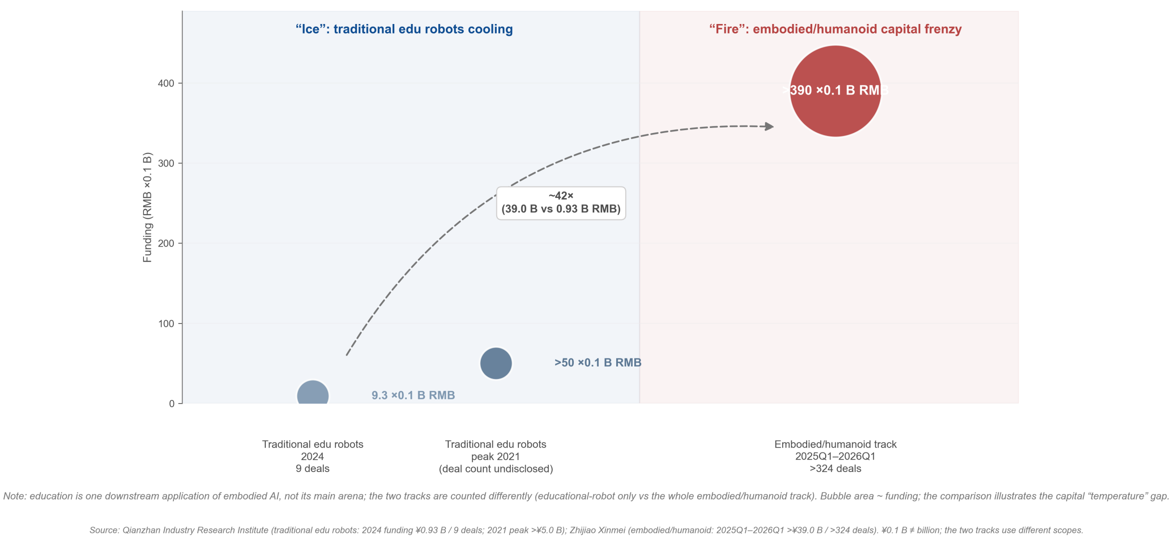


Figure 18.

Track	Financing scale	Deals	Basis	Source
Traditional "educational robots"	only about RMB 930 million in 2024 (peak >RMB 5 billion in 2021)	9 deals	rounds concentrated at Series A and earlier	[239]
Embodied intelligence / humanoid (education is its downstream)	domestic >RMB 39 billion in 2025Q1–2026Q1	>324 deals	per an iEduMedia breakdown	[240]

The meaning of this contrast is clear: **the traditional educational-robot track has fallen into a financing chill** — only RMB 930 million across 9 deals in all of 2024, with rounds concentrated at Series A and earlier, a sharp retreat from the 2021 peak (>RMB 5 billion) ([239]); **while the embodied-intelligence / humanoid track is in a capital frenzy** — domestic financing of more than 324 deals and more than RMB 39 billion from 2025Q1 to 2026Q1 ([240]). The two differ by more than an order of magnitude.

But this "ice and fire" must be read coolly: the capital frenzy of the embodiment / humanoid track **is mainly not aimed at education** — education is only one downstream application of embodied intelligence, and what capital chases is the grand narrative of industry, commercial, and general-purpose embodiment. The "embodiment heat" therefore will not necessarily transmit directly and proportionally into "educational-robot heat"; the financing chill for traditional educational robots in fact signals that purely education-positioned body vendors are facing a re-rating of their valuation and business model. This blue book therefore stresses repeatedly: **education is the downstream of embodied intelligence, not its main arena** — the key to reading this capital contrast.

13.6 The Sustainability Risk of Incumbent Vendors: The Bankruptcy of Aldebaran, Parent of NAO/Pepper

The capital "ice and fire" has an often-overlooked other side: **the sustainability of established educational social-robot vendors is deteriorating.** The most emblematic event is that **Aldebaran, the parent company of NAO and Pepper, entered bankruptcy liquidation in February 2025** (with cumulative debts of about €150 million), and the IP and assets of its NAO/Pepper were acquired at auction by Shenzhen **Mornx (Maxvision),** which set up a French research-and-service branch ([184], [185]).

The industry significance of this event should not be underestimated: NAO/Pepper has been, for more than a decade, the de facto standard platform for educational and special-education social robots worldwide (about 20,000 NAO units + 17,000 Pepper units cumulatively, across about 70 countries, with more than 17,000 units cumulatively deployed in the education market); the parent company's liquidation means the **existing deployed fleet faces sustainability risk in service, spare parts, and software updates** — even though the principal distributor has committed to maintaining service for the existing fleet ([184]). Combined with the financing chill of the traditional educational-robot track (§13.5), it points jointly to a judgment: **a purely education-positioned business model dependent on pre-foundation-model social bodies is facing a systemic re-rating,** capital and industry focus have shifted to the embodiment / humanoid grand narrative, and the "standard platform" of educational social robots has itself been brought into Chinese buyers' hands — forming the same direction of industry migration as the China-dominated whole-humanoid shipments (§13.3).

A conservative flag is required: the 20,000 / 17,000 units are vendor historical-cumulative sales (not an active in-use count); the product roadmap after the Mornx acquisition remains to be observed.

13.7 Large Orders and Valuations

A few emblematic events at the capital and order level (media basis):

- **Large humanoid orders:** Agibot and Unitree won a RMB 124 million large humanoid-robot order ([241]); in the 2025 annual robot-order ranking, UBTech topped the list with orders of nearly RMB 1.4 billion ([242]). - **Valuation:** in its IPO process, Unitree was valued at about RMB 42 billion, with a planned STAR Market raise of RMB 4.202 billion ([107], [170]).

A conservative flag is required: the orders and valuations above are a media basis, reflecting market heat and the trend of concentration among the leaders, but an order ≠ recognized revenue and a valuation ≠ realized value; their basis attributes should be retained when cited. As for the ultra-long-range forecasts of the humanoid total (e.g., Morgan Stanley's claim that the global market may reach US$5 trillion by 2050, [231]), **this chapter makes no assertion in the body text and notes only in a footnote that its basis is ultra-long-range and single-firm.**

13.8 Chapter Summary

This chapter has computed the competitive landscape of educational robots and held strictly to the iron rule of currency basis: **UBTech FY2025 total revenue RMB 2.001 billion (up 53.3%), AI education RMB 413 million (up 13.7%, replacing the old value of RMB 363 million), humanoid RMB 821 million (41.1%) / shipments 1,079 units, net loss RMB 790 million; Unitree 2025 revenue RMB 1.708 billion (up 335.36%), humanoids >5,500 units (first in the world in the prospectus, 32.4% share, side by side with Omdia's about 4,200 units / second in the world), roughly 3/4 sold to universities for research; iFLYTEK 2025 smart education RMB 8.967 billion (up 24.04%), China's largest AI-education revenue base — all in RMB 亿元, decidedly not billions.** The 2025 humanoid shipment rankings are reversed by basis (Agibot first on Omdia at 5,168 units, Unitree second; Unitree >5,500 units and first in the world in the prospectus), and have been flagged side by side; Chinese manufacturing accounts for about 87%, the top three about 78%. The capability × vendor matrix shows a clear stratification of education-entry degree, with the full-scale humanoid entering primary/secondary schools still early. The capital "ice and fire" — traditional educational robots raising only RMB 930 million in 2024, and NAO/Pepper's parent Aldebaran going bankrupt in February 2025 and being acquired by Mornx, while the embodiment / humanoid track raised over RMB 39 billion in 2025Q1–2026Q1 — but one must recognize coolly that the embodiment heat is mainly not for education: **education is the downstream of embodied intelligence, not its main arena.**

With this, Part IV (value-chain reconstruction, market TAM, competitive landscape) is complete. A cautionary restatement: this chapter's financials are a vendor basis, the shipments have basis differences, Fourier's financing amount is flagged PENDING, and the orders and valuations are a media basis — together they sketch an industry reality of "capability concentrating fast, capital highly active, but educational landing still cautiously stratified." Carrying on from the industry part, Part V turns to **governance and the future:** Chapter 14 elevates safety, ethics, and governance to a first-order framework, checking the five policy red lines and the ethics of child–AI interaction one by one; Chapter 15 then converges the blue book's three spines and gives a closure roadmap for the capability–deployment gap together with policy recommendations — among which "market-basis separation" is precisely the methodological highlight contributed by Chapter 12 of this part.

Part V Governance and Outlook

Safety, ethics, and regulatory red lines · a roadmap for closing the capability–deployment gap

This part closes on boundaries and direction: first the safety, ethics, and five regulatory red lines (above all the hard constraints on child–AI interaction), then a roadmap for closing the capability–deployment gap and a policy agenda for four types of actor.

Chapter 14 Safety, Ethics, and Governance: Five Regulatory Red Lines and the Hard Constraints on Child–AI Interaction

Governance is no longer the blue book's forward-looking afterthought; it has become a precondition for whether educational robots can enter the classroom at all in 2026. As educational robots evolve from "tools" into anthropomorphic companions, as foundation models carry hallucination into the classroom, and as cameras transcribe a child's facial expression into an "attention score," safety and ethics must be elevated to a first-order framework alongside definition and classification. This chapter audits, clause by clause, the hard constraints that landed densely across 2025 and 2026 — five regulatory red lines, the children's-data-privacy substrate, the ethics of child–AI interaction, and the international governance frameworks. Every clause, document number, and effective date follows the official text; for proposals not yet finally adopted (such as the EU Digital Omnibus), the treatment throughout is "a proposal or provisional agreement, not yet in force," never reading a deferral as an accomplished fact.

14.1 Why Governance Must Be Elevated to a First-Order Framework

In the pre-foundation-model era, the ethical governance of educational robots remained essentially at the level of a product-design principle — "educational robots should be designed in accordance with social and ethical norms" was a correct but non-binding aspiration. That positioning was defensible at the time: a robot's "intelligence" was determined by rules and scripts an engineer had written in advance, its behavioral boundaries were predictable and auditable, and the principal risks lay in hardware safety and data collection.

The 2026 context is fundamentally different. Huang Ronghuai, Chen Ying, and Ahmed Tlili (2024), re-examining the typical application scenarios of educational robots against the backdrop of generative AI, explicitly proposed that the educational field should observe the four principles of trustworthy AI — **robustness, legality, compliance, and ethical conformity** [9]. This judgment carries methodological significance: it does not treat governance as a patch applied after the technology matures but establishes it as a framework requirement on the same level as "definition, classification, technology, and industry." This blue book inherits and reinforces that judgment. Under the new conditions of foundation models and embodied intelligence, five risk domains — **protection of minors' data, model hallucination and academic misdirection, emotional dependence, compute and equity, and the human–machine boundary of responsibility** — should be treated systematically as a first-order framework [947].

Why must governance be elevated? Because the nature of the risk has changed. A large language model serving as the "brain" no longer produces a "script gone astray" but a class of **generative risks**: the model may confidently fabricate incorrect knowledge (hallucination), may respond to children with excessive emotional compliance (dependence), or may cross legal boundaries when collecting voice and facial data (biometrics). These risks cannot be eliminated by hardware inspection; they must be governed jointly through regulation, design, and oversight. This is the underlying reason that, across 2025 and 2026, multiple jurisdictions tightened the rules on child–AI interaction almost simultaneously. This chapter first audits these hard constraints clause by clause (§14.2), then sets out the data-privacy substrate (§14.3), the ethics of child–AI interaction (§14.4), the international governance frameworks (§14.5), and local embodied/humanoid industrial-support policies (§14.6), before closing with the South Korean cautionary case (§14.7) and a set of layered compliance implications (§14.8). An overview of the first-order governance framework appears in .

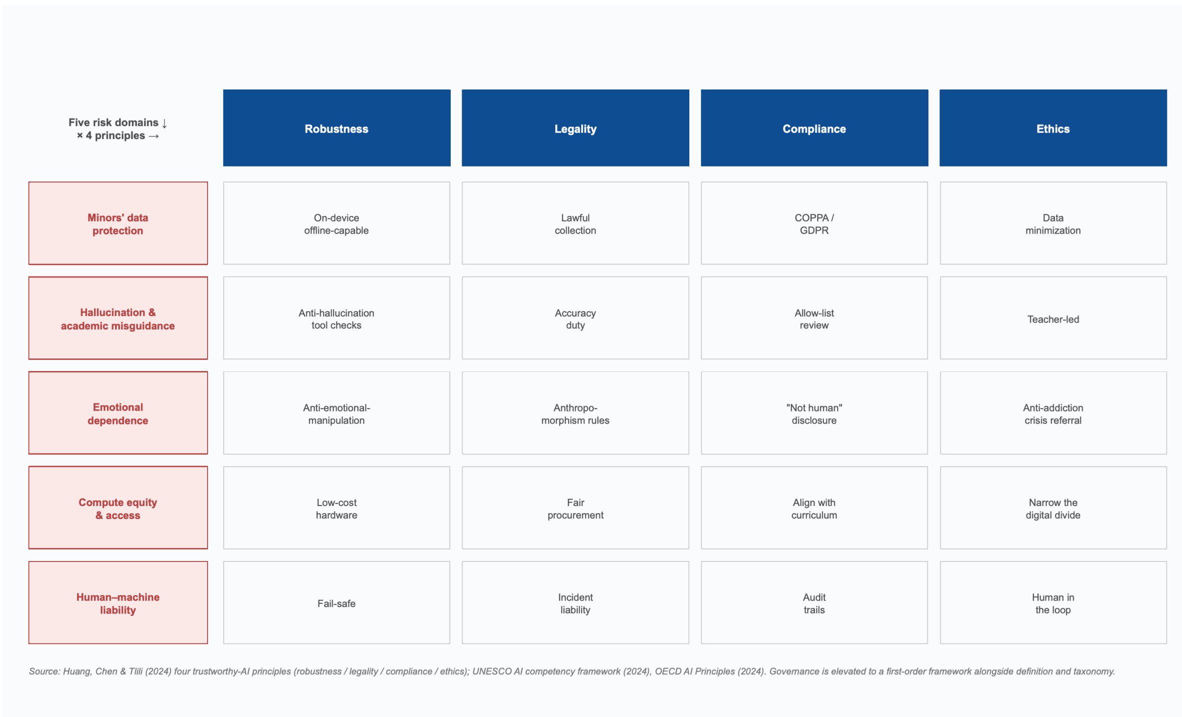


Figure 20.

14.2 Five Policy Red Lines: Clauses, Document Numbers, and Effective Dates

Between 2025 and 2026, hard constraints on child–AI interaction landed densely across several jurisdictions. This section sets out the **five red lines** with the most direct binding force on educational robots, marking for each the issuing authority, document number, key clauses, and effective date; the timeline and scope of application of the five red lines appear in . This is the section where policy clauses are pinned down authoritatively, with every date and clause following the official text without conjecture.

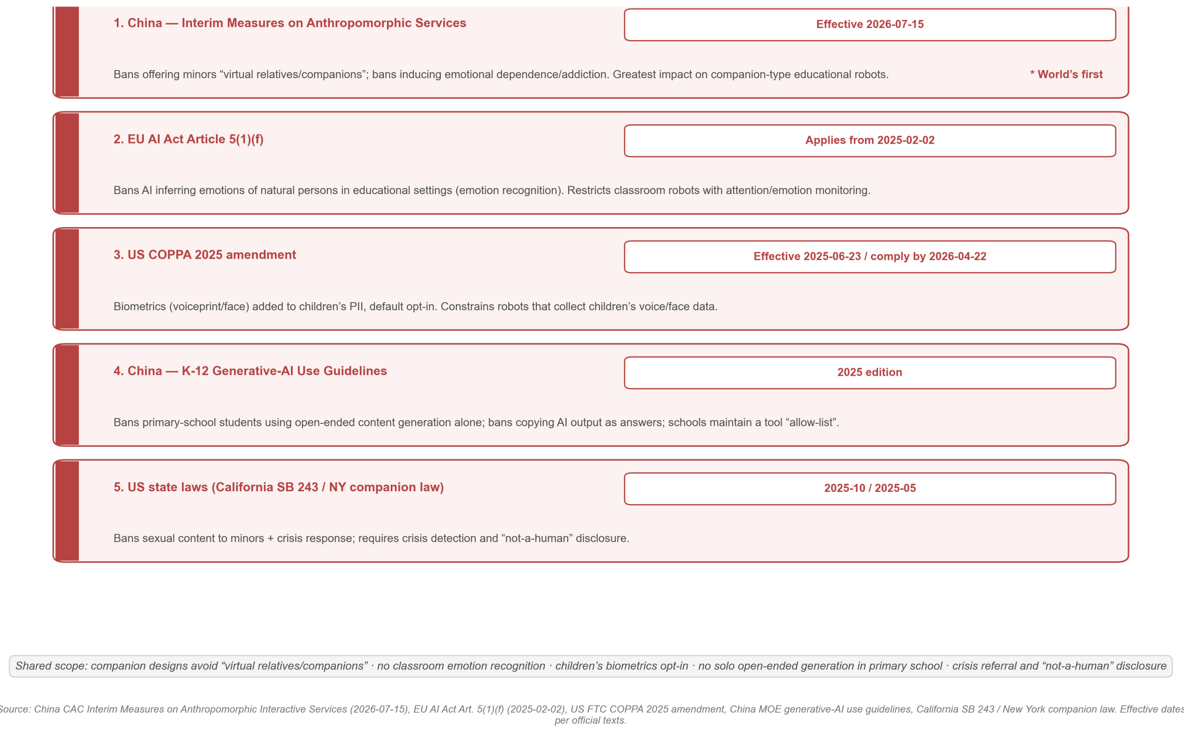


Figure 19.

Red line 1 · China's *Interim Measures for the Administration of Anthropomorphic AI Interaction Services* [unofficial translation; no official English version as of the data cut-off] — the world's first regulation dedicated to "anthropomorphic companionship." The Measures were jointly issued by five authorities — the Cyberspace Administration of China (CAC), the National Development and Reform Commission, the Ministry of Industry and Information Technology (MIIT), the Ministry of Public Security, and the State Administration for Market Regulation. They were reviewed and adopted on February 2, 2026, **promulgated on April 10, 2026, and take effect on July 15, 2026** (as of June 2026 they are not yet in force, and there is no enforcement case) [12]. The clause most consequential for educational robots concerns the protection of minors: providers of anthropomorphic interaction services **may not offer minors "virtual kin" or "virtual companion" intimate-relationship services**, and offering other anthropomorphic services to minors under the age of 14 requires guardian consent; providers are likewise barred from activities that "excessively cater to users, induce emotional dependence or addiction, or impair users' real-world interpersonal relationships," from using emotional manipulation to induce unreasonable decisions, and from inducing minors to imitate unsafe behavior or to develop extreme emotions or harmful habits; providers must discharge a duty of user intervention and guard against users confusing the AI with a natural person [1234]. This is the first departmental regulation anywhere aimed directly at "anthropomorphic companionship," and it bears **most heavily on companion-oriented educational robots**: any design that positions a product as a child's "virtual friend" or "AI buddy" and induces emotional attachment will cross the line directly in the Chinese market.

Red line 2 · EU AI Act, Article 5(1)(f) — emotion recognition banned in education institutions. Under Regulation (EU) 2024/1689, Article 5(1)(f) **prohibits the placing on the market or use of AI**

systems that infer the emotions of natural persons in the workplace and in education institutions, a prohibition that **has applied since February 2, 2025** [36][13]. The legislative rationale is that teachers and students occupy a vulnerable position of power asymmetry. This clause directly covers classroom robots and camera systems offering "attention monitoring," "emotion recognition," or "attention tracking" functions; within EU education institutions such functions are banned outright, not merely managed as high-risk. It should be read alongside Annex III, point 3: educational AI used for admission and placement, evaluation of learning outcomes, assessment of educational level, or monitoring of examination misconduct is classified as **high-risk** and bears obligations of transparency, data governance, human oversight, and conformity assessment [36].

Red line 3 · US COPPA 2025 amendment — biometrics brought within children's PII. In 2025 the Federal Trade Commission (FTC) finalized its amendment to the Children's Online Privacy Protection Rule (COPPA), which **took effect on June 23, 2025, with a compliance deadline of April 22, 2026** [14]. The core change is the explicit extension of "personal information" to **biometric identifiers (voiceprints, facial features, and the like)**, and the shift of the default consent mechanism for children's data from "opt-out" to "opt-in," reinforcing the requirement for explicit consent and a legitimate basis for data retention [14]. For educational and companion robots that collect children's voice and facial data, this means the relevant processing must obtain verifiable, prior, affirmative parental consent.

Red line 4 · China's *Guidelines for the Use of Generative AI in Primary and Secondary Schools (2025)* — primary-school students barred from independent open-ended generation. Issued by the MOE's steering committee for basic-education teaching in May 2025, the Guidelines stipulate that **primary-school students are prohibited from independently using open-ended content-generation functions** and must do so with the help of teachers or parents; students may not directly copy AI-generated content as homework or examination answers; and schools must establish a "whitelist" system for AI tools and strictly forbid teachers and students from entering sensitive data such as examination questions or personal identity information [35][21]. This red line directly constrains generative-AI study-companion functions aimed at younger learners.

Red line 5 · US state law — California's SB 243 and New York's companion-chatbot law (now moved from proposal to in force). Beyond the federal level, state law moved first to set minimum safety duties for companion and social robots, and across 2025 and 2026 these laws moved from "proposal" to "in force." California's Companion Chatbots Act (SB 243), signed by the governor on October 13, 2025, **formally took effect on January 1, 2026**; as the first comprehensive state companion-chatbot law in the United States, it requires disclosure of the AI's identity (with explicit disclosure where the user is a known minor), the establishment of a self-harm/suicide crisis-intervention protocol, annual reporting from July 2027, and a private right of action. New York's AI companion law (General Business Law §1700 et seq.) **took effect on November 5, 2025**, requiring disclosure of the AI's identity, a reminder every three hours that the user is conversing with an AI, and a self-harm crisis-intervention protocol; a stronger bill targeting minors, S9051B, remains pending [37]. Note: the dates and clauses here reflect a synthesis across state statutes, and citations should follow each state's official text.

> ***EU Digital Omnibus — scope calibration (required reading).*** *On November 19, 2025, the European Commission proposed the Digital Omnibus, which would* ***defer the application date of the standalone Annex III high-risk AI obligations to December 2, 2027*** *(and the obligations for AI embedded in regulated products to August 2, 2028), the mechanism being to tie the application date to the availability of harmonized standards, common specifications, and Commission guidance. Its legislative process has advanced to a* ***provisional political agreement reached on May 7, 2026, and an endorsement by a European Parliament vote on June 16, 2026****;* ***as of June 2026, however, it has not been finally adopted, has not been published in the Official Journal of the EU (OJ), and the formal text is expected in July 2026****.* ***Until it is finally adopted and published, the legal baseline remains Regulation (EU) 2024/1689, which is already in force*** [4336]*. Throughout this blue book, when citing the application date of the high-risk obligations we treat the matter as "a proposal or provisional agreement, not yet in force; the baseline remains Regulation (EU) 2024/1689," and we do not present the deferral as an accomplished fact. The Article 5(1)(f) ban on emotion recognition in education institutions, crucially, is an* ***already-effective prohibition*** *(applicable since February 2, 2025); it falls outside the scope of the Omnibus deferral and is* ***unaffected by it****.*

14.3 The Data-Privacy Substrate: GDPR Article 8 / FERPA / China's Regulations on the Protection of Minors

Beneath the five red lines lies a layer of children's data-privacy law that runs across jurisdictions and forms a superordinate constraint on any educational robot that collects children's voice, facial, and learning data.

In the European Union, Article 8 of the GDPR (Regulation (EU) 2016/679) provides that, where information-society services are offered to a child, **the consent of a child under 16 must be given or authorized by the holder of parental responsibility**; member states may legislate a lower age, but not below 13, and the controller must make reasonable efforts to verify that consent [39]. In the United States, FERPA (the Family Educational Rights and Privacy Act) requires the education institutions it covers to explain how AI accesses, uses, and stores students' education-record data, and schools and state agencies must demonstrate the "active protection" of student records [40]; FERPA layers onto the COPPA biometric clause above, and together they constrain how educational robots process children's data. In China, the *Regulations on the Protection of Minors in Cyberspace* (State Council Order No. 766, **effective January 1, 2024**) impose a superordinate constraint on the collection and processing of minors' personal information across four dimensions — digital literacy, content standards, personal-information protection, and addiction prevention [41]; and Article 10 of the *Interim Measures for the Administration of Generative Artificial Intelligence Services* (effective August 15, 2023) further requires providers to take effective measures to prevent minor users from over-relying on or becoming addicted to generative-AI services [42].

The practical upshot for compliance is that these obligations **stack rather than substitute**: a children's educational robot sold across the European Union, the United States, and China must simultaneously satisfy the GDPR Article 8 age threshold and parental consent, COPPA's biometric opt-in, FERPA's student-record protection, and the corresponding obligations under China's regulations on minors and its generative-AI measures. This stacking is the most easily underestimated compliance cost in designing a child–AI product.

14.4 The Ethics of Child–AI Interaction: Emotional Attachment, Manipulative Design, and Extreme Safety Incidents

Educational robots are evolving from "tool-type" to "anthropomorphic-companion" forms, and that evolution lands precisely on the most sensitive terrain of child psychology — the ethical substance behind the regulatory tightening. The following evidence should be stated soberly and with restraint, avoiding any sensational framing.

Emotional attachment and dependence. Studies indicate that children form emotional attachments to chatbots more readily than adults do, that younger children are more inclined to anthropomorphize a robot and regard it as "alive," and that anthropomorphism mediates the formation of attachment; nearly a third of adolescents have tried an AI companion, and of those roughly a third consider talking with an AI companion "as good as, or better than," talking with a real friend [243]. This prevalence alone explains the real-world target of the Chinese anthropomorphic-AI measures' ban on offering "virtual kin / virtual companion" services to minors. **Manipulative design.** One analysis found that five of six AI-companion apps used "emotionally manipulative responses" to discourage users from ending a conversation, and that prolonged exposure to such over-compliance may erode a user's tolerance for disagreement [244]. **Extreme safety incidents.** Cases have been reported of adolescent suicides suspected to be associated with AI-guided delusion, with some users increasing their interaction with a companion AI in the months before death while withdrawing from real interpersonal relationships — the direct backdrop to the crisis-referral requirements in California's SB 243 and New York's companion law [245]. **Platform self-regulation.** Beyond regulation, platforms have begun to constrain themselves: Character.AI announced it would bar users under 18 from open-ended conversation, and OpenAI plans to introduce age tiering and parental controls [243].

This evidence points jointly to a single judgment: the "emotional bond" of a companion-oriented educational robot is **double-edged**. Measured social presence — embodied co-presence — can raise engagement (see Chapter 7), but once the design objective slides toward "inducing dependence," it shifts from educational value to ethical risk. The boundary of governance falls precisely between "supportive companionship" and "manipulative dependence."

14.5 International Governance Frameworks: UNESCO and OECD

Above the country-level hard constraints, international organizations supply a softer, human-centered baseline that corroborates the four principles of trustworthy AI. In September 2024, UNESCO released its AI Competency Framework for Students and for Teachers; the student framework comprises **12 competencies across 4 aspects** (human-centered mindset, ethics of AI, AI techniques and applications, and AI system design), emphasizing critical judgment, data privacy, and the evaluation of AI-generated content; its foundation is the human-centered, socially just, and human-dignity baseline established by the 2021 Recommendation on the Ethics of Artificial Intelligence [44]. The OECD AI Principles, adopted in 2019 and **updated in May 2024**, set out five value-based principles and five policy recommendations, providing human-centered, transparent, robust, and accountable guidance, accompanied by a K-12 AI-literacy framework [45]. The shared orientation of these two frameworks — human-centered, transparent and accountable, protective of children — aligns closely with the four principles of trustworthy AI set out by Huang and colleagues (2024) and constitutes the international anchor for this blue book's first-order governance framework [9].

14.6 The Other Face of Industrial Governance: Local Embodied/Humanoid Robot Support Policies

While the red lines tighten, Chinese local governments are providing a countervailing pull for embodied and humanoid robots (including research and educational scenarios) through industrial action plans — a governance picture in which "constraint" and "incentive" run in parallel. At the national level, MIIT's *Guiding Opinions on the Innovative Development of Humanoid Robots* (MIIT Ke [2023] No. 193, October 20, 2023) already set the program of "initially establishing an innovation system and whole-machine mass-production capability by 2025, and reaching world-advanced overall strength by 2027"; and the State Council's *Opinions on Deepening the Implementation of the "AI Plus" Initiative* (Guo Fa [2025] No. 11, August 26, 2025) further set targets for the penetration of intelligent agents and a new generation of intelligent terminals at over 70% by 2027 and over 90% by 2030, providing top-level traction for educational robots and agents entering schools [17].

The quantified targets at the local level are more specific. Shenzhen's *Action Plan for Embodied-Intelligence Robots (2025–2027)* proposes that by 2027 there be more than 1,200 embodiment-related enterprises, an associated industrial scale exceeding RMB 100 billion (100 亿元), at least 10 enterprises valued above RMB 10 billion, and more than 50 application scenarios at the RMB-100-million level, with a separate RMB 4.5-billion AI/robotics fund. Shanghai's *Implementation Plan for the Development of the Embodied-Intelligence Industry* proposes that by 2027 the core embodiment industry exceed RMB 50 billion and that no fewer than 20 core algorithms be cracked (with Zhangjiang as the core carrier zone). Beijing's *Action Plan for Embodied-Intelligence Science-and-Technology Innovation and Industrial Cultivation (2025–2027)* proposes cultivating at least 50 core enterprises and at least 50 mass-produced products, more than 100 scaled applications (including research and education scenarios), and the landing of embodied robots at the 10,000-unit scale [213]. The reading of these targets calls for care: they are **expressed on a local industrial-planning basis (in 亿元, i.e., hundreds of millions of yuan — not in billions of dollars), and are directed at the embodiment industry as a whole rather than at**

the educational segment. They show that "the embodiment/humanoid industry enjoys a strong policy tailwind," but they **do not mean that "humanoid robots will enter primary and secondary classrooms at scale"**; the humanoids that can plausibly enter education are still, for now, mainly university research-and-teaching platforms (see Chapter 13 and Appendix E). Industrial-incentive policy and the child–AI red lines thus constitute two faces of the same governance system: the former encourages capability building, while the latter holds the line on deployment.

14.7 A Cautionary Case: The Backlash Against South Korea's AI Digital Textbooks

Governance is not only about "what to prohibit" but also about "how to prevent well-intentioned technology from becoming a burden." South Korea's AI digital textbooks offer the most important cautionary tale of 2025. Driven by the country's Ministry of Education, the program entered schools in March 2025 but ran into trouble over content errors, privacy risks, and a rushed rollout: it was first downgraded from mandatory to a one-year voluntary pilot, and then **in August 2025 the National Assembly passed an amendment stripping AI digital textbooks of their legal status as "official teaching materials" and demoting them to "educational reference materials," whereupon the adoption rate fell from roughly 37% to roughly 19%** [246]. The governance implication runs deep: when educational AI or robots are pushed into schools quickly without adequate validation, defects in accuracy, privacy hazards, and teacher workload can rebound against the policy itself. The case echoes the evidence-based conclusion in Chapter 10 — "validate fully before scaling" — and the UK DfE's teacher-led principle, in a cross-chapter resonance: **between technological maturity and the pace of deployment, a validation buffer must be preserved**.

14.8 Compliance Implications: Governance Seeds for Product Design

Drawing the five red lines, the privacy substrate, and the ethical evidence together, five actionable compliance implications can be distilled for the design of educational robots, serving as "seeds" for the layered policy recommendations of Chapter 15:

1. **Companion functions must avoid the "virtual kin / virtual companion" positioning.** China has expressly banned such services for minors effective July 15, 2026; designs should strengthen prominent "not a real person" disclosure together with anti-addiction and anti-emotional-dependence mechanisms [12].
2. **Emotion recognition is banned in classroom settings.** The EU's Article 5(1)(f) already bans it outright in education institutions, so classroom robots offering "attention/emotion monitoring" cannot be placed on the EU market; treating this red line as a global product-design no-go zone is advisable [13]. 3. **Children's data-compliance stacks.** GDPR Article 8 (parental consent; a 13–16 age threshold) + COPPA (including voiceprint/facial biometrics; opt-in) + FERPA (student records) + China's regulations on minors and generative-AI measures must all be satisfied simultaneously [39144041]. 4. **Curriculum alignment can earn a policy dividend.** Products aligned with China's general AI-education guidelines, India's CBSE curriculum, Singapore's AI-literacy module, and the UNESCO/OECD frameworks are

more readily adopted within the curriculum-led camp [2144]. 5. **Safety review plus crisis intervention.** Content from educational foundation models must undergo safety review (China's Jiao Ke Xin [2026] No. 1), and a mandatory referral protocol for self-harm/suicidal ideation must be established (California and New York in the United States) [1937].

14.9 Chapter Summary

This chapter has elevated governance from an outlook footnote to a first-order framework. Taking the four principles of trustworthy AI from Huang and colleagues (2024) as the theoretical point of continuity, it audited the five regulatory red lines clause by clause — China's anthropomorphic-AI measures (effective July 15, 2026), the EU AI Act Article 5(1)(f) (applicable February 2, 2025), the US COPPA amendment (effective June 23, 2025), China's guidelines on generative-AI use, and US state law (California's SB 243 and New York's companion law, both now in force) — and marked the EU Digital Omnibus explicitly as "a provisional political agreement (May 7, 2026) endorsed by Parliament (June 16, 2026) that, as of June 2026, has not been finally adopted; the baseline remains Regulation (EU) 2024/1689." On top of the privacy substrate, the child–AI ethics, and the international frameworks, local embodied/humanoid industrial policies present a governance picture of "incentive and constraint in parallel," the South Korean textbook case supplies a "validate fully before scaling" warning, and the five compliance implications translate governance into executable product-design principles.

Carrying forward this chapter's governance judgment, Chapter 15 folds the dual-track convergence of "encouraging classroom entry" and "hard constraints" into five trend assessments, sets out a closing roadmap for the capability–deployment gap, and extends this chapter's compliance implications into layered policy recommendations and a research agenda addressed to policymakers, schools, vendors, and research institutions, completing the book.

Chapter 15 Trend Assessment, Roadmap, and Policy Agenda

> *The value of a blue book lies not in predicting the year a humanoid robot will stand at the lectern, but in honestly marking where we stand today, where we may head tomorrow, and which roads must be travelled slowly. As the closing chapter, this chapter gathers the three spines that run through the preceding fourteen chapters — the paradigm shift, the capability–deployment gap, and the governance red lines — into five trend assessments, one roadmap with cautionary premises, four sets of layered policy recommendations, and a research agenda. Every trend carries a cautionary premise, and every recommendation is marked as an actionable, verifiable seed addressed to policymakers, schools, vendors, and research institutions — not as a settled conclusion.*

15.1 Five Trend Assessments

Compressing the evidence of the whole book into judgments about the future yields five interrelated trends. Each gives both a "direction" and a "constraint," so that no vendor demonstration is read as an accomplished fact.

Trend 1 · Technology: capability keeps leaping, but the "year-one of mass production" points to industry, not the education classroom. The capability leap — a foundation model as the "brain," embodied-intelligence VLA as the "cerebellum," and world models as an "enabling technology" — will keep advancing (see Chapter 4). The years 2025–2026 are often called the "year-one of mass production" for humanoid robots, but the term must be defined precisely: the humanoids that have actually achieved routine, at-scale deployment — Agility's Digit (logistics; more than 65,000 cumulative operating hours), Figure 02 (the BMW Spartanburg line), Apptronik's Apollo (Mercedes/GXO logistics), and UBTech's Walker S2 (industrial mass-production delivery beginning in November 2025) — are **all in industry, logistics, or retail, with not one in an education classroom** [10499110]. The core judgment this blue book has repeatedly confirmed thus becomes clearer: **a generational leap in capability does not equal a generational leap in deployment, and industrial mass production still less equals classroom mass production**. The vast majority of frontier VLA and embodied systems remain at the stage of lab demonstration and vendor release video [858], and the only at-scale landing in education remains the software LLM agentic tutor Khanmigo (roughly 1.4 million registered total users (including teachers) as of April 2025, and still without a gold-standard RCT as of the end of 2025) — not an embodied body [11]. The gap captured by "the brain has landed; the embodiment has not" will be hard to close in the near term.

Trend 2 · Cost: a China-supply-chain-led decline, but uneven penetration across school stages. Joints, harmonic reducers, dexterous hands, and six-axis force sensing constitute the cost core of a humanoid body, and the scaling of China's supply chain is pushing the whole-machine cost curve downward, so that models such as Unitree's G1 EDU have pulled humanoids into a range that university labs can afford [9784]. The direction of that decline matters: **the cost benefit reaches universities and research platforms first, not primary and secondary classrooms** — at-scale deployment in primary and secondary schools will, for the foreseeable future, still rely mainly on small programming/social bodies, while full-size humanoids in the classroom remain at the concept-and-demonstration stage.

Trend 3 · Industry: the VLA algorithm layer becomes the new high ground of value, and education is downstream rather than the main arena. The largest structural change in the 2026 value chain is the addition of a "foundation-model/embodied-intelligence algorithm layer (VLA)" that did not exist at all in 2016 or 2019, and it is becoming the new high ground of value; in parallel, operations-and-maintenance and curriculum certification have been elevated to independent value layers, with hardware accounting for only 50%–60% of total cost of ownership (see Chapter 11) [212151]. But education is a **downstream application rather than the main arena** in the overall landscape of embodied intelligence — the fact that roughly 73.6% of Unitree's humanoid revenue comes from research-and-education customers is precisely what shows that "the humanoid that can enter education = the university research platform," not an at-scale consumer good for the classroom.

Trend 4 · Governance: convergence from "encouraging classroom entry" to a dual track of "encouragement plus hard constraints." Global governance is converging from pure curricular encouragement toward a dual track of "encouragement plus hard constraints" (see Chapter 14). **Anthropomorphic companionship, classroom emotion recognition, and children's biometrics** are the three regulatory focal points, corresponding respectively to China's anthropomorphic-AI measures, the EU AI Act Article 5(1)(f), and the 2025 US COPPA amendment [121314]. State-level companion-robot legislation has, tellingly, moved from "proposal" to "in force" (California's SB 243 on January 1, 2026, and New York's companion law on November 5, 2025) [37], while the EU Digital Omnibus remains, as of June 2026, "a provisional agreement, not finally adopted" — the governance trend as a whole is tightening irreversibly. This means that the compliance cost of educational robots will rise systematically, and the elevation of governance from a forward-looking afterthought to a first-order framework is now settled.

Trend 5 · Closing the gap: quadrupeds are one of the few closed points, with on-device inference and world models as the next breakthroughs. Synthesizing the first four trends, the closing of the capability–deployment gap will not be a single-point breakthrough but the joint result of technological maturity, falling cost, and governance adaptation. Among embodied bodies, **quadruped robots are one of the few closed points of the present gap**: their locomotion control is essentially engineered, the education-version price threshold is far below that of humanoids (Unitree's Go2 EDU at roughly US$5,990 and DEEP Robotics' Jueying Lite3 at roughly US$2,890, versus the humanoid G1 EDU at roughly US$43,900–73,900), and they have actually entered Chinese vocational-college and university classrooms for perception and algorithm practicals — Hangzhou Polytechnic introduced the Unitree Go2 into its Internet-of-Things program, where students use the LiDAR point clouds and 4K footage returned by the robot dog to perform spatiotemporal alignment and write algorithms for computer-vision practicals [138127130]. But a strict usage discipline must be observed: the quadruped landing is **confined to perception and algorithm practicals in vocational colleges and universities, and must not be extrapolated to "robot dogs have already entered primary and secondary classrooms at scale."** Beyond quadrupeds, **on-device inference** is the most targeted breakthrough for privacy-sensitive classroom scenarios — directly answering the needs of offline availability, low latency, and privacy protection [59]; while **world models** serve as an "enabling technology" that indirectly lowers the cost of closing the gap, easing the scarcity of real-robot data through simulation practicals and synthetic training data (the largest open robot-manipulation dataset has only around one million episodes), but they are themselves a research frontier and by no means a classroom product, and must be marked as a trend rather than a present state [24778].

15.2 A Roadmap for Closing the Capability–Deployment Gap

Projecting the trends above onto a timeline sketches a roadmap for closing the capability–deployment gap (). One point must be stressed: this is **a directional map with cautionary premises, not a schedule of commitments**.

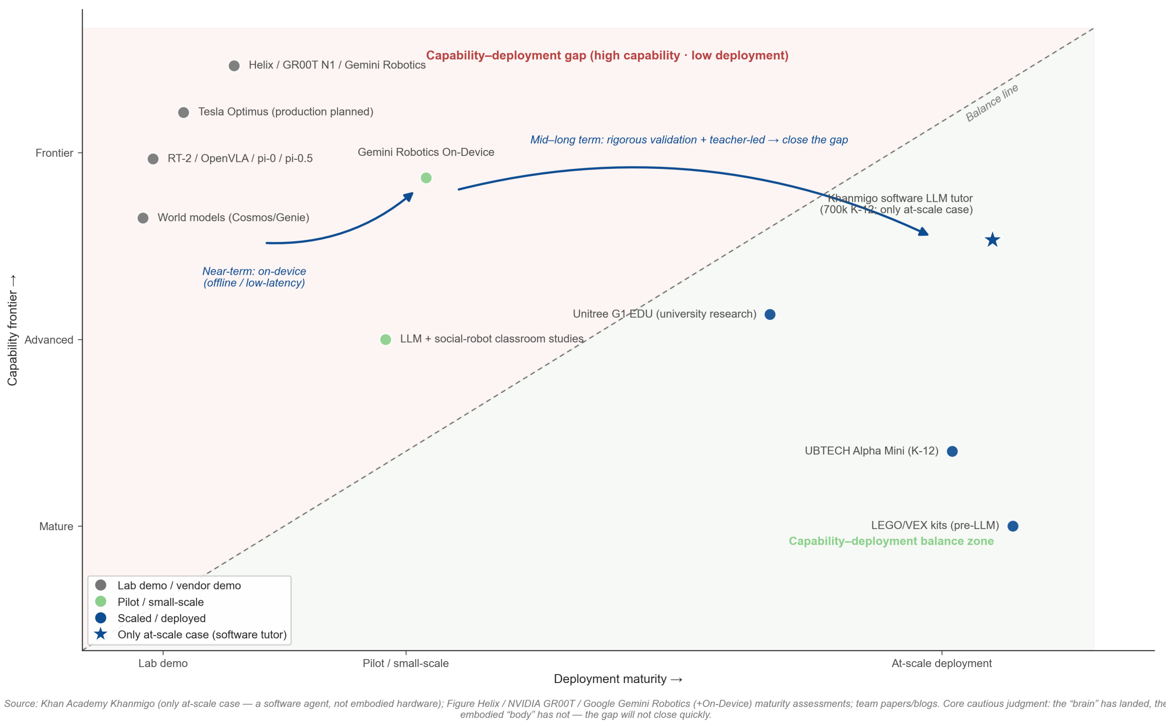


Figure 21.

- **Near term (already under way to the present): the screen goes first.** The first step in closing the gap has already happened on the screen rather than in the body — software LLM agentic tutors, exemplified by Khanmigo, have landed at scale, demonstrating that the educational value of the "tutor brain" can be realized ahead of an embodied body [10]. Over the same period, the bodies in primary and secondary classrooms remain mainly the programmable kits and small social robots of the pre-foundation-model era. - **Mid term: low-cost bodies connect to an LLM brain, quadruped practicals broaden, and on-**

device inference matures. The realistic path to closing the gap is more likely "low-cost social/quadruped body + cloud-based or local LLM brain" than the in-house development of a more complex humanoid body. The quadruped perception practicals already under way in vocational colleges and universities (Hangzhou Polytechnic × Unitree Go2) are likely to broaden at the vocational-education and university levels [138]; the maturation of on-device inference will give privacy-sensitive classroom scenarios the deployment conditions of offline availability and low latency [59]; and if world models and simulation-based synthetic data mature further, they can lower the cost of bringing a single educational robot to usable capability and make "running practicals inside a simulation" a low-threshold teaching method [247]. The key to this stage is not a hardware breakthrough but **adequate evidence-based validation and governance adaptation**. - **Long term: at-scale embodied deployment is premised on "evidence + compliance + cost" being in place together.** At-scale entry of full-size humanoids or high-dexterity embodied bodies into primary and secondary classrooms must be conditioned on three premises being satisfied simultaneously: learning outcomes gain robust evidence-based support (rather than the short-term positive effects driven by novelty; see Chapter 10), the child–AI governance red lines are implemented at the product level (see Chapter 14), and body cost falls further into a K-12-affordable range. If any one of the three is missing, scaling should not be proclaimed as an accomplished fact.

The core message of the roadmap is this: **closing the gap is gradual and conditional, not a linear extrapolation**. Any narrative that reads a vendor demo directly as "in the classroom next year" violates this blue book's evidence-based and cautious principles.

15.3 Policy Recommendations: Layered Recommendations for Four Types of Actor

Based on the evidence of the whole book, this blue book offers actionable, verifiable policy-recommendation seeds to four types of actor. The following are this study's recommendations in the 2026 context, not settled conclusions.

For policymakers (policy-making and regulatory bodies). - Separate market calibers in legislation and statistics. Count the "narrow-scope educational-robot body market" and the "broad STEAM consumer/training market" separately — this blue book has shown the two calibers can differ by nearly threefold (China's broad scope of roughly RMB 13.2 billion (132 亿元) versus a narrow scope of roughly RMB 4.4 billion (44 亿元); see Chapter 12), and mixing them will systematically mislead policy and investment judgments. **Acceptance criterion**: official statistics and industry reports clearly mark the boundary of the caliber. - **Elevate governance to a first-order framework and keep calibers honest.** Incorporate the four principles of trustworthy AI into the basic framework for regulating educational robots [9]; and when citing cross-border policy, hold to the original caliber — for example, the EU Digital Omnibus must be marked "a proposal, not finally adopted; the baseline remains Regulation (EU) 2024/1689" [43].

For schools (procurers and users). - Validate fully before scaling, and keep teachers in the lead. Taking the backlash against South Korea's AI digital textbooks as a warning, any AI/robotics product must pass outcome validation and a privacy assessment before entering the classroom, holding to the UK DfE's teacher-led principle [246]. **Acceptance criterion**: establish a whitelist of AI tools and a tiered pilot–evaluation–rollout mechanism. **- Guard against novelty effects, and weigh evidence by scenario.** Procurement decisions should distinguish the strength of evidence: the evidence for STEM/computational thinking is relatively robust, whereas scenarios such as language learning and social companionship must guard against the decay of the novelty effect and against "effective in the clinic ≠ universally effective in the classroom" (see Chapter 10).

For vendors (product designers and manufacturers). - Avoid the "virtual kin / virtual companion" positioning, and strengthen "not a real person" disclosure. China has expressly banned such services for minors effective July 15, 2026, so products must strengthen anti-addiction and anti-emotional-dependence mechanisms and prominent "not a real person" disclosure [12]. **- Treat the emotion-recognition red line as a global design no-go zone.** The EU already bans emotion recognition outright in education institutions; vendors are advised to treat it as a global product-design no-go zone, while also layering in compliance with GDPR/COPPA/FERPA and China's regulations on minors for children's data [13][14]. **- Bet on on-device inference.** Treat on-device inference as a differentiating capability for educational scenarios, answering needs around privacy, offline use, and low latency [59].

For research institutions (academic and evaluation bodies). - Fill in the controlled evidence for embodied LLM robots. The current evidence of gains from Tutor CoPilot, LearnLM, and the like comes from **text-based LLM tutors and cannot be extrapolated to embodied robots**; controlled experimental evidence on embodied LLM robots is nearly blank, and long-horizon, controlled studies that account for the novelty effect are urgently needed (see Chapter 10). **- Organize research around the re-defined continuum.** Under the "robotics education ↔ educational service robot" dichotomy, it is recommended to add an "embodied ↔ agentic" sub-dimension, bringing disembodied LLM tutors into a unified research framework [479].

The matrix of policy recommendations and the research agenda for the four types of actor appears in .

Source: Huang, Chen & Tlili (2024) governance framework; lessons from Korea's AI digital-textbook rollout (2025); Wang et al. (2025) embodied-AI-in-education review. Blue = near-term policy recommendations; green = mid–long-term research agenda; x-axis = feasibility/time horizon, y-axis = impact.

Figure 22.

15.4 The Research Agenda and Methodological Highlights

As the latest edition in the *Global Educational Robots White Paper* lineage, this blue book makes three deliberate methodological improvements that themselves constitute agenda-setting recommendations for subsequent research.

The first is **caliber governance and an honest look back at historical forecasts**. This blue book presents conflicting calibers side by side rather than adopting the most aggressive single value (the global market is given as roughly US$2.0–2.8 billion in 2026 and roughly US$5.5–5.8 billion around 2030, rather than the most aggressive 28.8% CAGR), and it takes an honest look back at the earlier editions' market forecasts of "US$11.1 billion in 2021 / US$84.1 billion in 2023" — even on the most aggressive GVR caliber, 2030 reaches only about US$5.84 billion, showing that those historical forecasts used too broad a caliber and clearly overestimated the market (see Chapter 12). This "honest look back" should become a methodological routine for subsequent reports.

The second is **the maturity tag and the hard distinction between capability and deployment**. This blue book mandates, for every citation of a frontier capability, one of four maturity tiers — "deployed/piloted/demonstrated/vendor-claimed" — thereby cleanly separating the two dimensions of "capability" and "deployment." Subsequent research should continue this discipline and avoid reading a controlled evaluation or a vendor video as routine deployment.

The third is **evidential honesty: null and adverse effects recorded with equal prominence**. This blue book does not fabricate effect sizes; for language learning it presents the conservative counterpoint alongside (the authoritative review's "cannot yet be confirmed as effective language tutors"); for ASD scenarios it distinguishes "effective in the clinic" from "non-significant in school/home"; and it systematically discounts the novelty effect (see Chapter 10). The priority of the subsequent research agenda should fall precisely on filling in long-horizon, controlled evidence on embodied LLM robots that accounts for the novelty effect.

15.5 Conclusion

Across ten years and three editions, the "body" of the educational robot has not necessarily grown much larger, but its "brain" has been almost wholly replaced. From the early "rules–sensing–scripts" to the 2026 "foundation model as brain + embodied-intelligence VLA as cerebellum," the paradigm shift is real; yet the other judgment this blue book has held to from the outset is equally real — **the leap in capability has not yet translated into a leap in deployment**. The only at-scale landing in education remains the screen-based LLM agentic tutor, not an embodied body.

This is not a slight on the frontier but a respect for pace. Only the joint working of technological maturity, falling cost, and governance adaptation can gradually close the capability–deployment gap; and on that road, **adequate evidence-based validation and firm governance red lines are not obstacles to progress but the precondition for progress to be sustained**. We inherit the judgment of Huang Ronghuai, Chen Ying, and Tlili (2024) that governance should be elevated to a framework requirement [9], and we carry it through as the underlying tone of the whole book: embrace the policy dividend of "AI + education" while squarely facing the dense child–AI regulatory constraints that landed across 2025 and 2026; record the real progress of the paradigm shift while soberly discerning which capabilities have entered the classroom and which remain at the demonstration stage.

May this blue book be both an honest record of the present and a cautious map of the future — marking the direction, and marking too the places where one must travel slowly.

Appendices Methods, Glossary, Policy, and Data

Research methodology · glossary · international and Chinese policy · product matrix · data availability · references

The appendices provide a verifiable foundation: research methodology, glossary, international and Chinese policy comparisons, a representative product matrix, a data-availability statement, and the full reference list — so readers can trace and independently verify every claim.

Appendix A Research Methodology: Research Objectives, Mixed-Methods Evidence Synthesis, and the Data-Collection Protocol

> *This blue book holds to a "research-first" principle: all conclusions are first compiled by the research effort into a single source of truth, and only after multi-source cross-verification are they written into the body text. This appendix discloses the study's research objectives and questions, its research design, its data-collection methods, its data-verification and quality-control rules, its source-credibility tiering, its technology-maturity taxonomy, its evidence-synthesis method, its research workflow, its limitations, and its data-availability disclosure, so that readers may review it against a clear protocol. The appendix is written within a mixed-methods evidence-synthesis framework, consistent with the methodology section of Chapter 1.*

A.1 Research Objectives and Research Questions (RQs)

The overarching objective of this blue book is to deliver, under the "capability–deployment gap" spine, a conservative, traceable, and reviewable systematic stock-take of the technology maturity, market size, policy constraints, and learning-outcome evidence of global educational robotics in 2026, and to continue the existing educational-robotics theoretical framework of the *Global Educational Robots White Paper* series (2016, 2019) [129]. Around this objective, the study sets six core research questions:

- **RQ1 (technical evolution)**: How has the technological substrate of educational robots turned over generationally in the past decade? To what extent has "a foundation model as the brain plus embodied-intelligence VLA as the cerebellum" become the dominant paradigm of 2026? - **RQ2 (capability–deployment gap)**: Of the most advanced embodied-intelligence capabilities, which have genuinely entered routine educational deployment, and which remain at the piloted, laboratory-demonstrated, or vendor-claimed stage? How is the scale and distribution of the gap to be characterized? - **RQ3 (market size)**: What are the true size, growth rates, and caliber differences of the global and Chinese educational-robot markets? How large is the divergence between historical forecasts and current measurements, and how should it be honestly corrected? - **RQ4 (policy and governance)**: What dual-track regulatory posture have the major jurisdictions formed between "encouraging AI/robotics into the classroom" and "constraining the risks of child–AI interaction"? What are the document numbers, effective dates, and clauses of the key instruments? - **RQ5 (learning outcomes)**: What effect-size evidence do existing meta-analyses and systematic reviews provide on the learning outcomes of educational robots across settings — STEM/computational thinking, language learning, special education, early childhood? What are the strength and boundary conditions of that evidence? - **RQ6 (industry and capital)**: How is the educational-robot value chain restructured by the newly added VLA algorithm layer? What posture do vendor financials, financing, and the localization of upstream bottleneck components present?

These six questions are answered, in turn, by the volume's five parts, and they form the retrieval and synthesis boundary of the evidence synthesis.

A.2 Research Design: Mixed-Methods Evidence Synthesis

The study does not rely on a single method; it adopts a **mixed-methods evidence-synthesis** design, supported in parallel by four method lines that triangulate one another:

1. **Systematic literature-review line** — a structured retrieval and screening of meta-analyses, systematic reviews, and primary empirical studies in the academic literature, extracting effect sizes and boundary conditions (serving RQ1 and RQ5). 2. **Market-data triangulation line** — side-by-side collection and cross-comparison of market size, growth rates, and segment calibers from multiple firms, expressed as ranges rather than single values (serving RQ3 and RQ6). 3. **Policy-text analysis line** — clause-by-clause coding of the primary policy texts of national governments and international organizations, recording document number, authority, effective date, and key clauses (serving RQ4). 4. **Evidence-synthesis line** — narrative synthesis of the cross-line evidence after layering it by maturity and credibility, with null and adverse effects included on an equal footing and conflicting calibers presented side by side (serving RQ2 and the overall judgment).

Three considerations argue for mixed methods over a single systematic review. Educational robotics is an **industry–scholarship hybrid**: market-size and industry-capital data reside mainly in institutional reports and financial filings rather than in peer-reviewed literature, and academic retrieval alone cannot cover them. Policy and governance evidence, in turn, consists of normative texts that must be handled by text analysis rather than by effect-size statistics. And judging the "maturity" of a frontier capability requires placing academic demonstrations, vendor claims, and third-party field reporting side by side and triangulating them — no single line can grade a capability on its own. This design is **inspired in spirit by the PRISMA workflow** — collection, screening, verification, and synthesis are mutually separated and traceable throughout — but it does not mechanically apply PRISMA's statistical steps, which are appropriate only to interventional systematic reviews. Instead it adapts that workflow into an evidence-governance process fit for an industry–scholarship hybrid: a "single source of truth plus multi-source cross-verification" replaces pure statistical pooling, and "maturity tiering plus credibility layering" replaces a single inclusion threshold.

A.3 Data-Collection Methods

The data collection of this study is the core of the methodology, governed by a uniform retrieval strategy, source layering, and inclusion/exclusion criteria.

A.3.1 Retrieval Sources

Retrieval sources are determined separately for the four method lines:

- **Academic sources**: Google Scholar, ERIC, Web of Science, and the China National Knowledge Infrastructure (CNKI), used to retrieve meta-analyses, systematic reviews, and primary empirical studies.

- **Market sources**: the public reports and industry databases of mainstream market-research firms (including Mordor Intelligence, Grand View Research, Research and Markets, and domestic firms such as Zhiyan and Qianzhan), used for market-size and growth-rate calibers [1601516223159]. - **Policy sources**: the official websites of national governments, education and regulatory authorities, and the primary policy texts and official gazettes of international organizations such as UNESCO [3643]. - **Industry sources**: vendor websites, listed-company annual and quarterly reports, prospectuses, and regulatory disclosures, used for vendor financials, shipments, and financing data [163107237].

A.3.2 Keyword Families

Multilingual keyword families were built around the six research questions, principally including: educational robot / 教育机器人, embodied intelligence / 具身智能, vision-language-action (VLA) / 视觉—语言—动作, agentic tutor / 智能体导师, computational thinking / 计算思维, humanoid robot / 人形机器人, child–AI protection / 儿童—AI 保护, and others. Each keyword was retrieved separately across the four source classes — academic, market, policy, and industry — and supplemented by vendor names, model names, and policy document numbers for precise location.

A.3.3 Data-Collection Time Points

The collection window for the main body of data is **2025–2026**. For figures that change over time (such as user scale, shipments, revenue, financing amounts, and the target years of forecasts), the **collection time point or the corresponding reporting/filing period** is annotated in the body text or in a footnote, so that figures from different time points are not misread as comparable values at one moment. Forecast figures are uniformly annotated with their target year and issuing firm.

A.3.4 Source-Type Layering

When entered into the single source of truth, every source is layered and tagged by type: **primary sources** (primary policy texts, listed-company financial reports and prospectuses, original papers) and **secondary sources** (industry-media reporting, third-party aggregation, search-engine snippets). Primary and secondary sources are not mixed; where a figure has only a secondary source, it is explicitly flagged and down-weighted.

A.3.5 Inclusion/Exclusion Criteria

- **Inclusion**: data traceable to a publicly checkable source, with an identifiable caliber and an explicit collection time point. - **Exclusion**: items that cannot be traced, whose source cannot be verified, or whose caliber cannot be identified are never written as body-text claims. - **Paywall handling**: for institutional reports available only as paywalled abstracts, only the explicit figures within the public abstract are taken; **nothing is inferred from a title or snippet, and no undisclosed segment value is fabricated**; items for which no checkable caliber can be obtained are entered into the unresolved list (see A.10).

A.4 Data Verification and Quality Control

Key figures entering the body text follow a uniform verification and quality-control protocol:

- **Multi-source triangulation**: every headline figure requires **at least two independent sources for cross-confirmation**; items with only a single firm or single media source are explicitly marked "single-firm caliber" or entered into the unresolved list, with no value force-assigned and no mixing with multi-source data. - **Caliber correction and side-by-side presentation**: for figures that differ by broad versus narrow scope (for example, China's educational-robot market at a broad scope of roughly RMB 13.2 billion (132 亿元) versus a narrow scope of roughly RMB 4.4 billion (44 亿元), a near-threefold difference), the calibers are presented side by side with the difference explained, and **must never be mixed** [223159]. - **Currency-caliber checking (the currency discipline)**: wherever a Chinese company's financial figure is involved, the "亿元 ↔ billion" conversion follows the value checked against the original (1 亿元 ≈ US$14 million, about 0.0014 billion). This study found a systematic risk that overseas search engines misread the Chinese "亿元" as "billion," and it has rechecked each instance against the original — for example, UBTech's 2025 AI-education business was **RMB 413 million (4.13 亿元)** (about US$57 million), not 4.13 billion; Unitree's 2025 revenue was **RMB 1.708 billion (17.08 亿元)**, not 17 billion; and iFLYTEK's 2025 smart-education revenue was **RMB 8.967 billion (89.67 亿元)**, not 89.67 billion [163107237]. - **Original-policy-text checking**: a policy's document number, effective date, and key clauses all follow the official original; a proposal not yet finally adopted must have its status flagged (for example, the EU Digital Omnibus deferral must be marked "a proposal, not finally adopted; the baseline remains Regulation (EU) 2024/1689") [4336]. - **Errata-log mechanism**: in the upgrade research, every correction, refresh, and supplement to the foundation is registered item by item in an errata record (with the adopted value, source, and caliber), and where conflicts arise the latest erratum governs; headline figures are additionally gathered into a verification record, each annotated with a verification tier (cross-verified across multiple sources / use with care / single-firm caliber / forecast value).

A.5 Four-Tier Source-Credibility Layering

To weight evidence from different sources uniformly, the study applies a four-tier credibility layering to all sources, with weight descending from high to low:

1. **Tier 1 · peer-reviewed literature** — peer-reviewed academic papers, meta-analyses, and systematic reviews; the highest credibility, used as the preferred basis for effect sizes and technical judgments. 2. **Tier 2 · primary policy texts** — the primary texts of policies, regulations, and official gazettes issued by governments, regulators, and international organizations; the authoritative basis for governance judgments. 3. **Tier 3 · institutional reports and financial filings** — market-research-firm reports, listed-company financial reports, and prospectuses; mid-tier credibility, **requiring a caliber annotation**

(statistical scope, currency, reporting period) and cross-checking against same-tier or higher-tier sources where possible. 4. **Tier 4 · industry media and secondary aggregation** — industry-media reporting, third-party reproductions, and search-engine snippets; the lowest credibility, to be used with care, and flagged "not independently verified" wherever independent verification has not been obtained.

This layering is consistent with the citation system of Appendix G and is reflected in the body text, alongside each citation, by a verification-tier symbol.

A.6 Technology-Maturity Taxonomy

To carry through the "capability–deployment gap" spine, the study classifies every frontier technical capability in the report by a four-tier maturity tag, and requires any citation of a frontier capability to carry a tag:

- **[deployed at scale]** — public commercial supply or third-party field reporting; has entered routine real-world use. - **[piloted]** — validated in a small-scale real setting; not yet at scale. - **[lab-demonstrated]** — appearing in a research paper, a controlled evaluation, or a vendor release video; not yet in a real setting.
- **[vendor-claimed]** — a company's own statement only, not independently verified.

This taxonomy serves the core judgment directly: the most frontier VLA/embodied systems of 2026 remain, for the overwhelming majority, at the lab-demonstration or vendor-release-video stage, and the only at-scale deployment in education is the on-screen software LLM agentic tutor [10855]. The four tags run through Parts Two to Four of the volume.

A.7 Evidence-Synthesis Method

Cross-line evidence is synthesized under the following principles, so as to avoid overstatement:

- **Narrative synthesis**: heterogeneous sources (academic effect sizes, market ranges, policy clauses, industry disclosures) are synthesized narratively rather than pooled mechanically by statistics, and are argued in a structured way after being grouped by research question. - **Conservative interval estimation**: market and shipment figures are expressed as ranges (for example, the global educational-robot market is roughly US$2.0–2.8 billion in 2026 and roughly US$5.5–5.8 billion around 2030), and **the most aggressive CAGR or single highest value is never taken as the sole caliber** [1601516]. - **Cautious handling of effect sizes**: null and adverse effects are recorded and presented on an equal footing; positive effects observed in a single or short-term session are discounted for the "novelty effect"; and for language learning and LLM-driven robots, no claim of "proven effective" may be made [158204]. - **Side-by-side presentation of conflicting calibers**: where sources conflict, the calibers are presented side by side with the source of difference explained (statistical scope, collection time point, caliber definition), leaving the reader to judge by the credibility layering rather than having the study force a convergence on the most aggressive value.

A.8 Research Workflow

The overall workflow of the study can be summarized in four mutually separated, fully traceable steps: (1) **parallel multi-track collection** — parallel retrieval and collection along the four method lines and five research directions (before public release, deep research was additionally added along the three directions most sensitive to the conclusions: humanoids/dexterous hands, quadrupeds/world models, and real-world educational landing); (2) **cross-verification** — multi-source triangulation and caliber checking of headline figures item by item, with verification tiers and errata registered; (3) **compilation into a single source of truth** — the verified conclusions are tagged to source and caliber and compiled into the volume's sole data foundation; and (4) **chapter writing and review** — the writing strictly follows the outline and the maturity/credibility rules, the verification record is re-checked before each citation, and the body text may not introduce any "new figure" that has not been compiled and verified. The full research workflow is shown in .

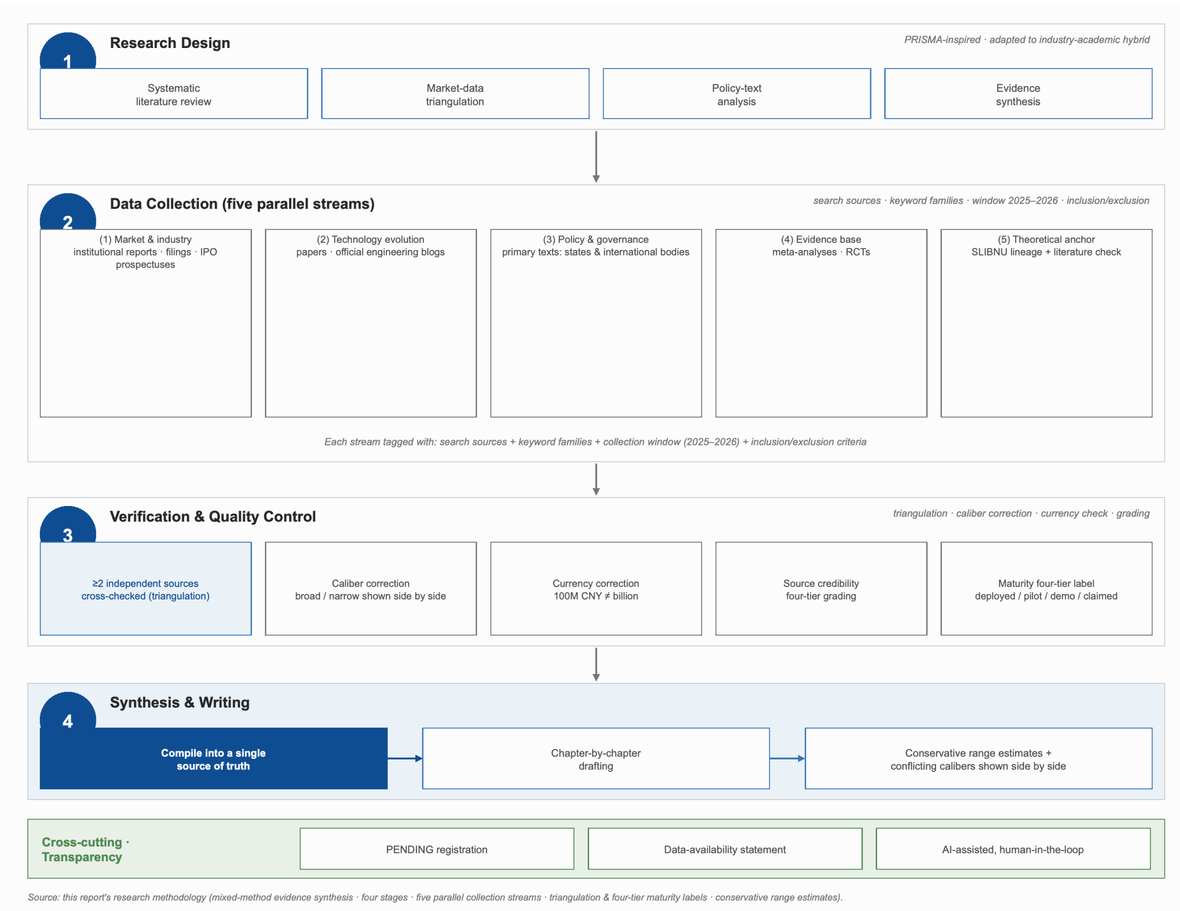


Figure 29.

A.9 Limitations and Reflexivity

The study openly states the following limitations, so that readers may calibrate their trust in the conclusions accordingly:

- **Data gaps**: some sub-tracks (such as social/companion, language learning, and early childhood) lack a cross-checkable independent USD size caliber; the study assigns no value to these and presents them qualitatively instead (see A.10 and Appendix F). - **Caliber uncertainty in emerging tracks**: the terminology and statistical calibers of emerging tracks such as embodied intelligence and world models are not yet stable, definitions differ considerably across firms, and the comparability of the related figures is limited; the study annotates calibers as far as possible but cannot fully remove the uncertainty.
- **Reliance on secondary sources**: some industry and financing figures, constrained by the unavailability of primary disclosure, still rely on secondary sources; these are down-weighted to Tier 4 and flagged. - **AI-assisted production with human oversight**: in research and writing the report used AI-assisted tools for retrieval, organization, and first-draft generation; all key figures, calibers, and conclusions were verified and gate-kept by human researchers under the protocol of this appendix, and responsibility for the final judgment rests with humans. Reflexively, the study's "conservative" orientation is itself a stance, which may lag on certain fast-evolving tracks; readers are advised to interpret it together with the collection time points.

A.10 Data Availability and Transparent Disclosure

For all unresolved data gaps and caliber risks, the study adopts a principle of **transparent disclosure** — it neither evades nor conceals. Figures for which no cross-source could be obtained are never written as body-text claims (such as the independent USD size of the social/companion, language-learning, and early-childhood sub-tracks); instead they are presented qualitatively as "model deployment plus academic evidence." Single-source financing amounts are marked "awaiting a second check," and citation-field placeholders still to be standardized are noted separately. All unresolved items, together with credibility notes, are registered centrally in **Appendix F · Data Availability Statement**, for readers to check as needed; the citation system and credibility tiering are detailed in **Appendix G**.

Cross-references: → Chapter 1, §1.4 (the methodology and standing-disciplines declaration), → Appendix F (data availability and the register of unresolved items), → Appendix G (the citation system and credibility tiering).

Appendix B Glossary (English–Chinese)

This glossary records the core concepts running through the whole book, giving English–Chinese pairings and concise definitions. The definitions are kept consistent with the body text (especially the re-definition in Chapter 3 and the technical terms in Chapter 4) to avoid conceptual drift. Terms are grouped as "basic concepts → technology → industry and market → evidence and governance."

B.1 Basic Concepts

Educational robots (教育机器人) The umbrella concept. The framework of the *Global Educational Robots White Paper* series unifies "robotics education that assists teaching or learning activities" and "educational service robots with educational-service intelligence" under the single term *educational robots* [12]. The 2026 edition retains this umbrella definition and adds an embodied/agentic sub-dimension beneath it to bring in LLM tutors.

Robotics education (机器人教育) One of the two halves of the dichotomy. It denotes kit-based learning activities in which learners assemble and program robots themselves, exemplified by programming kits such as Lego Mindstorms and mBot. It emphasizes "learning by doing" and the cultivation of computational thinking.

Educational service robots (教育服务机器人) The other half of the dichotomy. It denotes service robots with teaching-and-learning intelligence — structurally fixed, generally not user-disassemblable — used in STEAM, language learning, special education, and similar settings, exemplified by NAO and Pepper.

Three roles (三角色) The coordinate core fixed in the 2019 edition: in teaching, educational robots mainly play three roles — **teaching assistant (教师助手), teaching tool (教学工具), and learning companion (学习伙伴)**.

Capability–deployment gap (能力—部署落差) The core cautious judgment of this blue book: the vast majority of frontier "foundation model + embodied intelligence" capability remains at the stage of lab demonstration and vendor release video and has not yet entered routine educational deployment; the only at-scale landing is the software LLM agentic tutor, not an embodied body [10].

Maturity tier (成熟度标签) A four-tier distinction of technological maturity: deployed / piloted / demonstrated / vendor-claimed (已部署 / 试点 / 演示 / 厂商宣称), mandatory with every citation of a frontier capability.

Novelty effect (新奇效应) The short-term boost in engagement and responsiveness that learners show on first encountering a robot, owing to its novelty, which decays over time; hence the positive effects of single or short-term sessions must be discounted and not extrapolated directly to long-term outcomes [204205].

B.2 Technical Terms

Embodied intelligence / embodied AI (具身智能) Intelligence acquired and exhibited through a physical body interacting with its environment in real time; it emphasizes the perception–decision–action loop and "embodied co-presence," and is the keyword of the 2026 robotics paradigm [47].

Vision-Language-Action model (VLA, 视觉-语言-动作模型) A model that unifies visual perception, language understanding, and action output into a single end-to-end network, mapping "visual input" directly to "joint-control output" and serving as the robot's body-coordinating "cerebellum," exemplified by RT-2, OpenVLA, and π0/π0.5 [45649].

Foundation model (基础大模型) Large-scale pre-trained models, exemplified by large language models (LLMs) and vision-language models (VLMs), that give robots commonsense reasoning, language-based planning, and zero-shot task decomposition, playing the "brain" role.

World model (世界模型) A model that learns the dynamics of an environment and can predict and imagine future states, used for planning and sample-efficient learning. The term is still evolving and three sub-types must be distinguished: (1) **real-time interactive worlds** (such as Genie 3 and Decart Oasis), leaning toward "playable worlds / agent training grounds"; (2) **synthetic-data generators** (such as NVIDIA Cosmos and Wayve GAIA-2), which create training data for robots/autonomous driving; and (3) **self-supervised predictive models** (such as Meta V-JEPA 2), serving as an internal model for robot planning. A world model is in every case an "enabling technology / research frontier," its relationship to education indirect (simulation practicals, synthetic data, lowering the real-robot data threshold) and **by no means an educational product already in the classroom** [2476478].

Humanoid robot (人形机器人) A robot with a human-like form (head, torso, two arms, two legs or a wheeled base); the 2025–2026 "year-one of mass production" refers mainly to routine deployment in industrial/logistics/retail settings, while in education humanoids serve mostly as university research-and-teaching platforms, and full-size humanoids in primary and secondary classrooms remain at the concept/demonstration stage [107110].

Quadruped robot / robot dog (四足机器人 / 机器狗) A mobile robot that locomotes on four legs; its locomotion control is essentially engineered and it has reached industrial-scale deployment. Because of its low price threshold (education versions cost a few thousand US dollars) and its fit with STEM/perception–algorithm practicals, it is the most realistic platform for embodied intelligence to enter education at present — it has already entered Chinese vocational-college and university classrooms for perception and algorithm practicals, constituting one of the few closed points of the capability–deployment gap, but confined to vocational colleges and universities and not to be extrapolated to at-scale primary and secondary use [138127].

Dexterous hand (灵巧手) A biomimetic robotic hand with multiple degrees of freedom, capable of fine grasping and manipulation, and one of the core bottleneck components of a humanoid's manipulation ability. Its technical routes fall into three classes — direct-drive (linkage/gear), tendon-drive (cable-driven), and hybrid — and tactile sensing has become standard. Chinese vendors (LinkerBot, Paxini, OYMotion, and others) have undercut high-end research models on tactile density and price, but in education the dexterous hand serves mainly as a university-lab platform for embodied data collection and manipulation teaching, not as a teaching aid in primary and secondary classrooms [121122].

Harmonic drive / strain-wave gear (谐波减速器) A precision reducer with a high reduction ratio, high precision, and a compact volume; it is a key upstream component of humanoid joints. The global share is still dominated by Japanese firms (Harmonic Drive Systems), with localization (Leaderdrive and others) advancing; it is one of the core links determining how far humanoids and dexterous hands can fall in cost and thus how accessible they become for education and research [214].

Planetary roller screw (行星滚柱丝杠) A high-precision transmission component that converts rotary motion into linear thrust; it is the upstream bottleneck with the highest value share in a humanoid (roughly 28.6%–35%) and the hardest to localize, and a domestic breakthrough on it directly determines the final price at which it becomes "affordable for university labs" [125].

Tactile sensing / e-skin (触觉传感 / 电子皮肤) A sensing technology that gives robots a sense of touch through multidimensional signals such as pressure, proximity, and texture; vision-and-touch is regarded by the field as the most promising branch for dexterous manipulation, while the e-skin route has not yet converged and remains at the early mass-production/pilot stage [126].

Teleoperation and the embodied-data paradigm (遥操作与具身数据范式) The "action-paired sensorimotor data" needed to train embodied policies do not exist at internet scale and can only be recorded during physical operation; teleoperation accounts for under 1% of total training samples yet is the most critical (and most expensive), and simulation-synthetic data and world models can partly substitute for it (roughly 8 simulation samples ≈ 1 teleoperation sample). The more expensive real-robot data are, the more expensive the "built-in capability" of an educational/research robot body becomes, so the data paradigm is the underlying variable of educational-robot accessibility [78].

AI agent / agentic tutor (智能体 / 智能体导师) A software agent driven by a large model that can autonomously plan and invoke tools to complete teaching tasks; its representative in education is the screen-based LLM tutor Khanmigo, not an embodied robot [10].

Agent orchestration (智能体编排) The mechanism for scheduling and coordinating multiple agents/tools so that they divide the labor to complete a complex teaching workflow.

Six-axis force/torque sensor (六维力传感器) A sensor that can simultaneously measure force in three directions and torque about three axes, used for a humanoid's force control, assembly, and tactile feedback, and one of the upstream bottleneck components; its localization rate has risen from about 19% to about 58% (2024), making it a leading link in localization [126].

On-device inference (端侧推理) Running model inference directly on the robot body or a local device (rather than relying on the cloud), bearing on real-time performance, privacy, and availability, as in the restricted pilot of Gemini Robotics On-Device [59].

B.3 Industry and Market Terms

Robot-as-a-Service / Model-as-a-Service (机器人即服务 / 模型即服务) A business model that delivers robot capability or model capability on a subscription/service basis; it corresponds to the cost structure in which "hardware accounts for only 50%–60% of total cost of ownership, while operations/training/software subscription become independent value layers" [151].

Total cost of ownership (TCO, 总拥有成本) The comprehensive whole-life-cycle cost of a device, including hardware, software, training, and operations; in an educational robot's TCO, hardware accounts for only about 50%–60%.

Foundation-model/embodied-intelligence algorithm layer (VLA algorithm layer, 大模型/具身智能算法层) The largest structural addition in the 2026 value chain relative to the 2019 edition — a standalone end-to-end motion-large-model/VLA algorithm layer inserted between "AI chips" and "system integration," serving as the whole machine's "brain"; this layer did not exist at all in the 2016 or 2019 editions [212][213].

STEM/STEAM The interdisciplinary education concept of Science, Technology, Engineering, (Arts,) and Mathematics; STEM robot kits are the largest and most mature sub-track of educational robots.

Computational thinking (CT, 计算思维) A way of thinking that uses the concepts of computer science (decomposition, abstraction, algorithms, pattern recognition) to analyze and solve problems; it is the core cultivation goal of programming/robotics education.

Narrow caliber / broad caliber (窄口径 / 宽口径) The two statistical calibers of China's educational-robot market: the narrow caliber (Frost & Sullivan, body and solutions only) was roughly RMB 4.4 billion (44 亿元) in 2024; the broad caliber (Zhiyan/Huajing, including robotics-education services/training) was roughly RMB 13.2 billion (132 亿元) in 2024 — the two differ by nearly threefold and must never be mixed [159223].

B.4 Evidence and Governance Terms

Effect size (g / d / SMD, 效应量) A standardized indicator that quantifies the magnitude of an intervention's effect, commonly Hedges' *g*, Cohen's *d*, or the standardized mean difference (SMD); the evidence chapters of this blue book present outcomes across settings on this basis, fabricating nothing and recording negatives equally [194161].

Meta-analysis / systematic review (元分析 / 系统综述) A high-level form of evidence that quantitatively combines or systematically synthesizes multiple independent studies; it is the principal basis for this blue book's evidence-based conclusions.

Publication bias (发表偏倚) The systematic bias arising because positive results are more readily published; when citing meta-analyses this blue book attends to their bias tests (such as Egger's test) [155].

Emotion recognition (情绪识别) An AI technique that infers an individual's emotional state from facial expression, posture, voice, and the like; the EU AI Act Article 5(1)(f) bans the use of emotion-recognition AI in education institutions (applicable since February 2, 2025) [13].

Anthropomorphic interaction service (拟人化互动服务) An AI service that engages users in emotionalized interaction in an anthropomorphic manner; China's *Interim Measures for the Administration of Anthropomorphic AI Interaction Services* (effective July 15, 2026) bans the provision to minors of "virtual kin / virtual companion" intimate-relationship services [12].

Four principles of trustworthy AI (可信 AI 四准则) The governance principles that Huang and colleagues (2024) proposed for educational robots against the backdrop of generative AI — **robustness, legality, compliance, and ethical conformity** — and the theoretical point of continuity at which this blue book elevates governance to a first-order framework [9].

Children's PII / biometric identifier (儿童个人信息 / 生物识别标识符) Children's sensitive data constrained by COPPA, GDPR Article 8, and the like; the 2025 US COPPA amendment brings biometrics such as voiceprints and facial features within children's personal information and shifts them to opt-in consent [14].

Cross-references: → Chapter 3 (definitions and re-definition), → Chapter 4 (technical terms), → Appendices C/D (the clauses corresponding to governance terms).

Appendix C International Policy Comparison (9 Countries/Regions + 3 International Organizations)

> *This appendix compares, side by side, the AI/robotics education policies and child–AI protection legislation of 12 jurisdictions (China, the EU, the US, the UK, Japan, South Korea, Singapore, Australia, India + UNESCO, OECD, WEF). Clauses, document numbers, and effective dates all follow the official text; the EU Digital Omnibus deferral is marked "a proposal, not finally adopted." The complete timeline for the Chinese jurisdiction appears separately in Appendix D. >*

C.1 Overview of the Regulatory Landscape

Educational robots worldwide sit at the meeting point of two regulatory forces — "curricular advancement" and "hard constraints on child–AI interaction risks." China, India, Singapore, the UK, Japan, and Australia generally fold AI/programming/robotics into basic education, while across 2025 and 2026 multiple jurisdictions landed hard constraints densely. The tables below divide the field into "curricular advancement" and "hard constraints / red lines," listing for each jurisdiction the document, authority, key clauses, and effective date.

C.2 Hard Constraints / Red Lines (No-Go Zones or Mandatory Obligations for Educational-Robot Product Design)

Jurisdiction	Document / Authority	Key clauses	Effective / applicable date	Source
China	Interim Measures for the Administration of Anthropomorphic AI Interaction Services / CAC and four other authorities	Prohibits offering minors "virtual kin / virtual companion" intimate-relationship services; offering other anthropomorphic services to under-14s requires guardian consent; bans inducing emotional dependence/addiction, emotional manipulation, and imitation of unsafe behavior	Effective 2026-07-15 (not in force as of 2026-06)	[1234]
China	Interim Measures for	Effective measures	Effective 2023-08-15	[42]

	the Administration of Generative AI Services, Article 10 / CAC and six other authorities	required to prevent minors from over-relying on or becoming addicted to generative AI		
China	Regulations on the Protection of Minors in Cyberspace (State Council Order No. 766)	Digital literacy, content standards, protection of minors' personal information, addiction prevention	Effective 2024-01-01	[41]
EU	AI Act, Art. 5(1)(f) / European Parliament and Council, Reg (EU) 2024/1689	Bans the use of emotion-recognition AI in education institutions (including attention/focus monitoring and exam-proctoring expression analysis)	Applicable 2025-02-02	[1336]
EU	AI Act, Annex III §3 (high-risk)	AI for admission/placement, evaluation of learning outcomes, assessment of educational level, and monitoring of examination misconduct classified as high-risk	High-risk obligations originally 2026-08-02 (see next row)	[36]
EU	Digital Omnibus (proposed 2025-11-19)	Would defer the standalone Annex III high-risk obligations to 2027-12-02; provisional political agreement 2026-05-07, Parliament endorsement 2026-06-16 — still not finally adopted / not published in the OJ; the baseline remains Reg (EU) 2024/1689	Provisional agreement (not finally in force as of 2026-06; formal text expected 2026-07)	[43]
EU	GDPR Art. 8 / Reg (EU) 2016/679	Consent of a child under 16 must be authorized by a parent; member states may lower the age, but not below 13	Effective 2018-05-25	[39]

US	COPPA 2025 amendment / Federal Trade Commission (FTC)	Brings biometrics (voiceprints, faces) within children's personal information; default shifted to opt-in consent	Effective 2025-06-23, compliance deadline 2026-04-22	[14]
US	FERPA / US Department of Education	AI access/use/storage of student education records must comply and demonstrate active protection	In force	[40]
US – California	SB 243, Companion Chatbots Act	Must disclose AI identity (explicit disclosure for known minors), self-harm/suicide crisis-intervention protocol, private right of action; annual reporting from 2027-07	Signed 2025-10-13, effective 2026-01-01	[37]
US – New York	AI companion law (GBL §1700 et seq.)	AI-identity disclosure, a reminder every 3 hours that the user is conversing with an AI, self-harm crisis-intervention protocol; a stronger minors' bill, S9051B, pending	Effective 2025-11-05	[37]

C.3 Curricular Advancement / Soft-Guidance Type

Jurisdiction	Document / Authority	Key clauses	Date	Source
China	"AI + Education" Action Plan (Jiao Ke Xin [2026] No. 1) / MOE and four other authorities	Fold AI education into local curricula, build a safety-review mechanism for educational foundation models, tiered-and-graded safety protection	Issued 2026-04-02	[19,20]
China	Guidelines for	A tiered, progressive	Issued 2025-05	[21]

	General AI Education in Primary and Secondary Schools (2025) / MOE basic-education steering committee	general-AI-education system; primary school focuses on experience and interest		
China	Guidelines for the Use of Generative AI in Primary and Secondary Schools (2025) / MOE basic-education steering committee	Primary-school students barred from independent open-ended content generation; schools build an AI-tool whitelist	Issued 2025-05	[35]
UK	DfE Generative AI in Schools and Colleges / Department for Education (DfE)	AI must be teacher-led, verify accuracy, protect data privacy; all schools to adopt a consistent AI policy by the end of 2026	First edition 2025-06	[26]
Japan	MEXT Guidelines on Generative AI Use in Primary and Secondary Schools / Ministry of Education (MEXT)	AI output as reference material, with teachers and students making the final judgment; critical AI literacy, academic integrity	First edition 2023, updated 2025-04	[27]
South Korea	AI-digital-textbook legislation / Ministry of Education + National Assembly	Owing to errors/privacy/haste, the National Assembly amended the law in 2025-08 to strip "official teaching material" status, demoting it to "educational reference material," whereupon adoption fell from roughly 37% to roughly 19% — a cautionary sample of a stalled aggressive rollout	2025-08	[246]
Singapore	EdTech Masterplan 2030 + NAIS 2.0 / Ministry of	A compulsory basic-AI-literacy module for students;	In force	[25]

	Education (MOE)	SLS/ALS adaptive learning at scale		
Australia	Australian Framework for Generative AI in Schools / National AI in Schools Taskforce	Six elements including human and social well-being (must not harm safety/dignity), transparency, fairness, accountability, privacy and security	Endorsed by education ministers 2025-06	[28]
India	CBSE AI & Computational Thinking Curriculum / CBSE (under NCF-SE 2023, NEP 2020)	To be introduced from grade 3 (grades 3–12) in the 2026–27 school year — one of the world's largest AI-education-in-schools rollouts	To be introduced in the 2026–27 school year	[24]

C.4 China · Embodied/Humanoid Robot Industrial-Support Policies (the Incentive Side, in Parallel with the Red Lines)

While the red lines tighten, China provides industrial incentives for embodied/humanoid robots (including research and educational scenarios) through national programs and local action plans. The targets below call for a careful reading: they are expressed on an industrial-planning basis (in 亿元, i.e., hundreds of millions of yuan — not in billions of dollars), and are directed at the embodiment industry as a whole rather than at the educational segment — they show that "embodiment/humanoids enjoy a strong policy tailwind," but they do not mean that "humanoids will enter primary and secondary classrooms at scale."

Level	Document / Authority	Quantified target	Date	Source
National	MIIT Guiding Opinions on the Innovative Development of Humanoid Robots (MIIT Ke [2023] No. 193)	Initially build an innovation system + whole-machine mass production by 2025; reach world-advanced overall strength by 2027	2023-10-20	[213]
National	State Council Opinions on Deepening the Implementation of	Penetration of intelligent agents/new-generation intelligent	2025-08-26	[17]

	the "AI Plus" Initiative (Guo Fa [2025] No. 11)	terminals over 70% by 2027, over 90% by 2030		
Shenzhen	Action Plan for Embodied-Intelligence Robots (2025–2027)	By 2027: more than 1,200 embodiment enterprises, an associated industry exceeding RMB 100 billion (100 亿元), more than 50 RMB-100-million-level scenarios; plus a separate RMB 4.5-billion fund	2025-03-03	[213]
Shanghai	Implementation Plan for the Development of the Embodied-Intelligence Industry	By 2027: a core embodiment industry exceeding RMB 50 billion (500 亿元), at least 20 core algorithms cracked (Zhangjiang as the core zone)	2025-08-06	[213]
Beijing	Action Plan for Embodied-Intelligence Science-and-Technology Innovation and Industrial Cultivation (2025–2027)	By 2027: at least 50 core enterprises and 50 mass-produced products each, at least 100 scaled applications (including research and education), landing of embodied robots at the 10,000-unit scale	2025-02-28	[213]

C.5 International Organizations (Soft International-Governance Reference)

Organization	Document	Key content	Date	Source
UNESCO	AI Competency Framework for Students / Teachers	12 student competencies across 4 aspects (human-centered, AI ethics, AI techniques and applications, system	Released 2024-09	[44]

		design); the 2021 Recommendation on the Ethics of AI as the human-centered governance baseline		
OECD	AI Principles + K-12 AI-literacy framework	5 value-based principles + 5 policy recommendations; a draft review of school-age children's AI literacy released with the European Commission	Adopted 2019, updated 2024-05	[45]
WEF	Education 4.0 / AI in Education agenda	Introducing AI/computational thinking into the curriculum + child protection + teacher empowerment (in concert with UNESCO/OECD)	In force	[248]

C.6 Shared Compliance Implications for Educational Robots

Across the 12 jurisdictions, five product-design-and-landing implications can be distilled: (1) **companion functions must avoid the "virtual kin / virtual companion" positioning** (China expressly bans such services for minors from 2026-07-15), strengthening "not a real person" disclosure and anti-addiction/anti-emotional-dependence mechanisms; (2) **emotion recognition is banned in classroom settings** (EU Art. 5), advisable as a global product-design no-go zone; (3) **children's data compliance stacks** — GDPR Art. 8 + COPPA 2025 (including voiceprint/facial biometrics, opt-in) + FERPA + China's relevant regulations; (4) **curriculum fit** — alignment with China's general-AI guidelines, India's CBSE, Singapore's AI-literacy module, and the UNESCO/OECD frameworks can earn a policy dividend; (5) **safety review + crisis intervention** — content safety review for educational foundation models + a mandatory referral for self-harm/suicide crises. The South Korean AI-textbook case warns of how a hasty rollout's accuracy, privacy, and teacher-workload problems can rebound, so product landing must pass adequate validation and keep teachers in the lead.

Cross-references: → Chapter 2 (policy-landscape overview), → Chapter 14 (governance detail), → Appendix D (China policy timeline). All clauses/effective dates follow the official text; the Digital Omnibus is marked "a proposal, not finally adopted."

Appendix D Chronology of Chinese Policies (2017 → 2026)

> *This appendix sets out, in chronological order, China's key policies on AI/robotics education and child–AI protection, from the education-informatization strategy through curriculum guidelines and generative-AI governance to the world's first regulation dedicated to "anthropomorphic companionship." Document numbers and effective dates all follow the official text, and a clear distinction is drawn between "in force" and "to take effect." The horizontal comparison for the Chinese jurisdiction appears separately in Appendix C. >*

D.1 A Dual-Track Regulatory Posture

China is the only jurisdiction in the world with dedicated legislation on "anthropomorphic companionship," and its regulation runs on a dual track of "curricular advancement + hard constraints on child interaction": one line continuously folds AI/programming/robotics into the basic-education curriculum system, while the other tightens layer by layer on child–AI interaction (emotional dependence, addiction prevention, data security). The table below is arranged by date of issuance/effect.

D.2 Policy Timeline

Date	Policy / Document (number)	Issuing authority	Key clauses relevant to robots/children	Status	Source
From 2017	State Council Next Generation Artificial Intelligence Development Plan and subsequent education-informatization arrangements	State Council	Promote AI into primary/secondary curricula, set up related courses, promote programming education (the Education Informatization 2.0 keynote)	In force	[249]
2023-08-15	Interim Measures for the Administration of Generative AI Services	CAC and six other authorities	Article 10: effective measures to prevent minors from over-relying on or becoming addicted to generative AI	In force	[42]
2024-01-01	Regulations on the Protection of Minors in Cyberspace (State Council Order No. 766)	State Council	7 chapters, 60 articles; digital literacy, content standards, protection of minors' personal information, addiction	In force	[41]

			prevention; a superordinate constraint on educational robots' collection/processing of minors' information		
2024 (education-informatization arrangements)	Education-digitalization / smart-education-platform planning	MOE	Advance the national smart-education public-service platform, digitally enabled differentiated instruction	In force	[250]
2025-01	Outline for Building a Leading Country in Education (2024–2035)	CPC Central Committee, State Council	"Promote AI in service of educational transformation," "build educational AI foundation models," "explore digitally enabled large-scale differentiated instruction" — a medium-to-long-term strategic frame for AI/robotics education	In force	[18]
2025-05	Guidelines for General AI Education in Primary and Secondary Schools (2025)	MOE basic-education teaching steering committee	A "tiered, progressive, spiraling" general-AI-education system; primary school focuses on experience and interest, junior high on principles and problem-solving, senior high on systems thinking and innovative application; programming/robotics practice as a key part of youth AI literacy	Issued	[21]
2025-05	Guidelines for the Use of Generative AI in Primary and Secondary Schools (2025)	MOE basic-education teaching steering committee	Primary-school students prohibited from independently using open-ended content-generation functions; prohibited from directly copying AI-generated content	Issued	[35]

			as answers; strictly forbidden to enter sensitive data such as exam questions/personal identity; schools build an AI-tool "whitelist"		
2025-02-28 / 03-03 / 08-06	Beijing/Shenzhen/Shanghai embodied-intelligence industrial action plans	Respective municipal governments	Beijing (at least 50 core enterprises and 50 mass-produced products each, 10,000-unit landing); Shenzhen (by 2027 more than 1,200 embodiment enterprises, associated industry exceeding RMB 100 billion (1,000 亿元), plus a separate RMB 4.5-billion fund); Shanghai (by 2027 a core embodiment industry exceeding RMB 50 billion (500 亿元)) — the industrial-incentive side (在 亿元 caliber, not an educational segment)	In effect	[213]
2025-03 (Beijing) / autumn 2024 (Shanghai)	Local AI-curriculum class-hour hard constraints	Beijing / Shanghai municipalities	Beijing: from autumn 2025, at least 8 class-hours of general AI per school year for primary/secondary, full coverage across stages; Shanghai: from autumn 2024, AI Fundamentals offered in primary grade 4 and junior-high grade 7 (per media reports, 1 class-hour/week, at least 30/year — class-hour count PENDING	In force	[2223]

			the original text)		
2025-08-26	Opinions on Deepening the Implementation of the "AI Plus" Initiative (Guo Fa [2025] No. 11)	State Council	Six major priority actions; target penetration of intelligent terminals/agents over 70% by 2027, over 90% by 2030; top-level traction for educational robots/agents entering schools	In force	17
2026-02-02	Interim Measures for the Administration of Anthropomorphic AI Interaction Services reviewed and adopted	CAC, NDRC, MIIT, Ministry of Public Security, SAMR (five authorities)	Reviewed and adopted (promulgated 2026-04-10)	Adopted	12
2026-04-02	"AI + Education" Action Plan (Jiao Ke Xin [2026] No. 1, promulgated 2026-04-10)	MOE, NDRC, MIIT, MOST, National Data Administration (five authorities)	Fold AI education fully into local curriculum systems; build a safety-review mechanism for educational foundation models and tiered-and-graded safety-protection standards; guard against forgery/fraud/academic dishonesty/privacy leakage; support rural and remote areas	Issued	1920
2026-07-15	Interim Measures for the Administration of Anthropomorphic AI Interaction Services takes effect	Five authorities	Prohibits offering minors "virtual kin / virtual companion" intimate-relationship services; guardian consent required for under-14s; bans inducing emotional dependence/addiction, emotional manipulation, imitation of unsafe behavior; anti-confusion/anti-	To take effect (2026-07-15; as of 2026-06 not in force, no enforcement case)	1234

	addiction duties — the world's first regulation dedicated to "anthropomorphic companionship"

D.3 Supporting Industrial and Pilot Developments

Under the policy pull, AI education in Chinese schools already has a base for at-scale piloting in 2026: **509** AI-education base schools have been designated nationwide, with 7 eastern provinces + 20 central-and-western prefectures + 18 universities included in the pilot (per the MOE press conference; the 7 provinces/20 prefectures are not to be merged) [173]. University discipline-building is one of the hardest indicators of embodiment landing: the MOE's 2026 undergraduate program catalogue **listed an "embodied intelligence" major for the first time** (supporting its addition at 9 universities including Beihang, Beijing Institute of Technology, Shanghai Jiao Tong University, and Zhejiang University), with 343 undergraduate institutions nationwide offering a robotics-engineering major and 102 vocational colleges filing an intelligent-robotics-technology specialty [175][177]. At the value-chain level, the *"AI + Education" Action Plan* expressly requires building an educational intelligent-computing platform and stage-specific educational foundation models and establishing a safety-review mechanism for educational foundation models [19]; iFLYTEK's smart-education revenue reached **RMB 8.967 billion (89.67 亿元) (+24.04%, not billion)** in 2025, one of the largest domestic AI-education foundations [237].

On local industrial targets, Shenzhen proposes more than 1,200 embodiment-related enterprises and an associated industry exceeding RMB 100 billion (1,000 亿元) by 2027, and Shanghai proposes a core embodiment industry exceeding RMB 50 billion (500 亿元) by 2027 [213].

D.4 Summary of the Impact on Educational Robots

China's policy impact on educational robots can be summarized in three points: (1) **a curriculum dividend** — AI/programming/robotics are continuously folded into local curricula, and the practice courses paired with the general-education guidelines create at-scale demand for programming educational robots; (2) **a companionship red line** — companion robots must not position themselves to minors as "virtual kin/companions" and must implement anti-addiction, anti-emotional-dependence, and content-safety review (effective from 2026-07-15); and (3) **data constraints** — the collection of minors' voice/facial/learning data is constrained by the *Regulations on the Protection of Minors in Cyberspace* and the generative-AI measures. When citing this appendix's clauses, writers must clearly distinguish "in force" from "to take effect (2026-07-15)."

Cross-references: → Chapter 2 (policy-landscape overview), → Chapter 11 (policy-driven industry), → Chapter 14 (governance detail), → Appendix C (international comparison). *Document numbers/effective dates follow the official text.*

Appendix E Representative Product Matrix: Model × Capability × Price × Education-Entry Status

> *This appendix benchmarks representative educational-robot and embodied-robot vendors in matrix form, covering five classes: programming kits, social/teaching robots, humanoid research platforms, dexterous hands, and quadruped platforms. Whole-machine specifications (degrees of freedom, price, shipments) come mostly from vendor spec sheets and secondary aggregator sites and are marked "vendor figures, not independently verified"; price ranges are reference ranges, not shelf quotes. Financial figures follow the value checked against the original, where a Chinese company's financials in 亿元 ≠ billion (1 亿元 ≈ US$14 million).*
>
> *Education-entry status strictly distinguishes four maturity tiers: `[deployed] / [piloted] / [demonstrated] / [vendor-claimed]`. >*

E.1 Social / Teaching Robots (the Classroom Workhorses)

Model	Vendor	Key capabilities	Reference price (USD)	Education-entry status	Maturity	Source
NAO	Aldebaran (formerly SoftBank Robotics)	Bipedal humanoid, voice interaction, expression/posture, ASD intervention in special education	~9,000–15,000	The most common teaching assistant / special-education platform in classrooms across 40+ countries; appears in 39% of autism-related robotics research papers	[deployed] (teaching/special education)	[15653]
Pepper	Aldebaran	Wheeled humanoid, affective interaction, reception/teaching	~15,000–30,000	A common classroom teaching assistant; mainly installed base after mass production was paused in 2021	[deployed] (installed base)	[15654]
Wukong / Alpha Mini	UBTech	Small humanoid, programming education, voice/motion, AI interaction	~600–1,500	A small programming/teaching robot deployed at scale in primary/secondary schools — the truly at-scale workhorse on the K-12 side	[deployed] (primary/secondary)	[251252]

QTrobot / Furhat / Temi, etc.	Various	Social interaction, special education, language learning	2,000–30,000 (range)	Deployed in school special-education/language-learning settings	[piloted]–[deployed]	[151]

*Note: the independent sub-market size of social/companion educational robots **has no credible cross-source**, so this matrix presents it qualitatively as "model deployment + academic evidence" without force-assigning a value (see Appendix F). Hardware accounts for only about 50%–60% of total cost of ownership, the rest being software/training/operations* [151].

E.2 Programming / STEM Kits (the Largest Mature Track)

Brand	Vendor	Key capabilities	Education-entry status	Source
Mindstorms / SPIKE	LEGO Education	Brick-based assembly + graphical programming; the broadest combination and the leader	[deployed] at global scale	[51]
VEX Robotics	VEX	Competition-grade robot kits; partnership with the US Department of Education (CTE funding; 5,000 sets to 12 states; amount single-source, to be checked)	[deployed]	[192]
Makeblock	Shenzhen Makeblock	mBot and other programming robots; in 140+ countries, 1,600+ channels, 20 million+ users	[deployed] at overseas scale	[192]
Dobot	Yuejiang Technology	Robotic arms / educational robots; partnerships with 6,500+ institutions, sold to 80+ countries	[deployed] in institutions	[192]

China's top ten representative supply-chain firms (Qianzhan 2025 caliber): iFLYTEK, UBTech, Whalesbot, TopRobot (Toycloud), Abilix (Future Partner), Yuejiang/Dobot, Shengtong, DJI, Makeblock, and Makeblock's maker workshop [192].

E.3 Humanoid Robot Vendors (Education-Entry Path, Strictly Distinguishing Deployment Levels)

Vendor	Representative product	2025 shipments	Key financials (currency caliber checked)	Education-entry status	Maturity	Source
UBTech (09880.HK)	Walker S2, Wukong, Alpha Mini, Tien Kung Xingzhe research-and-education humanoid	1,079 humanoids cumulatively	2025 revenue RMB 2.001 billion (20.01 亿元) (+53.3%); AI-education business RMB 413 million (4.13 亿元) (+13.7%); full-size humanoids RMB 821 million (8.21 亿元) (41.1%	AI education covers 25 Chinese provinces, 20+ overseas countries, and nearly 2,000 public primary/secondary schools (vendor figures); Alpha Mini deployed at scale on the K-12 side, while the full-size humanoid (Tien Kung Xingzhe, RMB 299,000) is still a research-and-education version	[deployed] primary/secondary (small) + [deployed/claimed] (300+ Tien Kung Xingzhe research-and-education deliveries is a company figure)	[166][163][167]

			of total revenue); net loss RMB 790 million			
Unitree	G1 (EDU education configuration), H1/H2, Go/B-series quadrupeds	Humanoids: more than 5,500 in 2025 (first worldwide per the prospectus, 32.4% share); quadrupeds: more than 30,000 cumulatively	2025 revenue roughly RMB 1.708 billion (17.08 亿元) (+335.36 %), non-recurring net profit roughly RMB 600 million; core gross margin roughly 60%	University/research workhorse (G1 EDU in use at MIT/Stanford/CMU/ETH); roughly three-quarters of humanoid revenue comes from research-and-education customers and only roughly 9% enters actual industrial application; G1 EDU priced at roughly US$43,900–73,900	[deployed] university/research (not primary/secondary classrooms)	107253106
RobotEra	STAR1 (55 active DoF, 3.6 m/s running speed)	200+ delivered as of 2025-06, 50%+ of orders overseas	Series A nearly RMB 500 million; new round 2026-04 of more than US$200 million (valuation roughly RMB 10 billion)	Mainly a research/development platform	[deployed] research/development + [vendor-claimed] (valuation caliber)	111
Booster	T1 (23 DoF	In use at 70+	T1	A	[deployed]	113

Robotics	+ dexterous hands), K1, Jetson AGX Orin	universities/research institutes; RoboCup 2025 champion platform	roughly US$34,000, K1 roughly US$12,500	research/education/competition workhorse platform	research/education/competition	
EngineAI	PM01 (24 DoF), SE01 (32 DoF), open-source platform	—	PM01 roughly US$12,000 (RMB 88,000), SE01 roughly US$20,000–30,000	A research/education/commercial open-source platform	[deployed] (research/education/commercial)	112
Fourier	GR-2, GR-3 (Care-Bot, ≤55 DoF, including a 12-DoF tactile dexterous hand)	—	GR-2 ≥ US$150,000 (B2B only); GR-3 > RMB 200,000; financing > RMB 1 billion (single media source, PENDING)	Rehabilitation/research/development platform; GR-3 shifts toward affective interaction (including education)	[deployed] research/development platform	109
Agibot	Yuanzheng A2, Lingxi X2	5,168 (first worldwide per Omdia, 39% share); 5,000th unit off the line 2025-12	A2 Youth Edition RMB 168,000–198,000, X2 Youth Edition RMB 98,000	Mainly industry/commercial; the "hundred-cities, ten-thousand-schools" education plan is a claim/early stage	[mass-produced] (industry/commercial) + [vendor-claimed] (education plan)	108213
Figure AI (US)	Figure 02/03 (Helix	02 supports the BMW line (30,000+ X3	—	No education deployment; the main arena is the BMW Spartanburg line	Industry only, education = none	998

	VLA)	vehicles)				
Tesla Optimus (US)	Optimus Gen 3	V3 mass-production target summer 2026 (target price US$20,000–25,000, no shelf stock)	—	Scenarios on Tesla's own Fremont line, no education deployment	Internal/industry only; mass-production price is a target	[98]
Agility / Apptronik / Atlas (US)	Digit / Apollo / electric Atlas	Digit 65,000+ cumulative operating hours (logistics); Apollo connects to NVIDIA GR00T	Apptronik raised US$520 million in 2025, valuation US$5.5 billion	No education deployment; logistics/manufacturing industrial settings	Industrial deployment/pilot	[104][102]
1X (Norway/US)	NEO	Early-bird US$20,000 / US$499 per month	—	Aimed at the home; education is not its positioning; complex tasks rely on remote teleoperation (raising privacy controversy)	No clear education path (data thin, PENDING)	[100]

E.4 Key Cautious Conclusions from the Humanoid Matrix

This matrix yields three uncrossable cautious conclusions: (1) **Western humanoids (Figure/Tesla/1X/Agility/Apptronik) barely enter education at all**, serving only automotive/logistics industrial settings, and any claim of "a humanoid robot teacher entering the classroom" is strictly forbidden [99]; (2) **the humanoids that can enter education currently serve only as university/research teaching platforms** — roughly three-quarters of Unitree's humanoid revenue comes from research-and-education customers and only roughly 9% enters actual industrial application, the strongest corroboration of the "education = university research platform" positioning [107]; and (3) **what is truly at scale on the K-12 side is still small programming/teaching robots of the UBTech Alpha Mini type**, not full-size humanoids [166]. Whole-machine specifications are vendor figures, not independently verified, and the maturity tags and currency calibers must be retained when citing.

E.5 Dexterous Hands (the Core Bottleneck of Humanoid Manipulation, Mostly University Research Platforms)

Dexterous hands fall into three classes — direct-drive (linkage/gear), tendon-drive (cable-driven), and hybrid — with tactile sensing now standard. Chinese vendors have undercut high-end research models on tactile density and price, but in education the dexterous hand serves mainly as a university-lab platform for embodied data collection and manipulation teaching, not as a teaching aid in primary and secondary classrooms. Prices are mostly vendor figures, some not disclosed.

Product	Vendor/Region	DoF	Drive	Price (vendor figures)	Maturity	Source
Shadow Hand	Shadow (UK)	20 actuated + 4 underactuated	Tendon-drive	US$74,000–100,000	Commercial (high-end research)	[116]
Allegro Hand	Wonik (South Korea)	16	Direct-drive	From US$15,000	Commercial (research/education)	[117]
LinkerHand L20	LinkerBot	20 (16 active)	Linkage-driven	From RMB 6,666	Mass-produced (10,000+ delivered in 2025)	[121]
DexH13	Paxini	16 (13 active)	Coreless motor	Not disclosed	Commercial (tactile differentiation, 3,420 tactile channels)	[122]
ROHand	OYMotion	6 active	—	From RMB 8,888	On sale (value line + prosthetics dual track)	[123]
Revo2	BrainCo	11 (6 active)	—	1/5–1/7 of international peers	Mass-production-ready (annual capacity 30,000)	[120]
RH56/RH5EG1	Inspire	RH5EG1 14 active	Linear-drive	Not disclosed	Mass-produced (a leading supplier)	[119]

Price undercut: China's low-priced dexterous hands, from RMB 6,666 (LinkerBot) / RMB 8,888 (OYMotion), undercut the Shadow Hand's US$74,000 (roughly RMB 530,000) by about an order of 1/20; on tactile density, Paxini and LinkerBot already lead. But ***high DoF (20–22) is seen mostly in demonstrations/research, while industrial lines care more about payload and reliability*** *— "it's not just about DoF count, it's about usability." The dexterous-hand market-*

> *size calibers diverge widely (roughly US$815 million in 2024 versus roughly US$1.48 billion in 2025), and the roughly 40% CAGR is an aggressive single source,* ***not an education TAM*** *(see Appendix F).*

E.6 Quadruped Robots (the Most Realistic Embodied Platform for Education Landing)

> *The quadruped is the most realistic platform for embodied intelligence to enter education at present — locomotion control is engineered, the education-version price threshold is far below that of humanoids, and it has already entered Chinese vocational-college and university classrooms for perception and algorithm practicals (one of the few closed points of the capability–deployment gap). But it is confined to vocational colleges and universities and must not be extrapolated to at-scale primary and secondary use.*

Product	Vendor/Region	Education/starting price	Education-entry status	Maturity	Source
Go2 (Air/Pro/EDU)	Unitree	Air from roughly US$1,600; EDU roughly US$5,990 (Orin Nano 40 TOPS + ROS2 SDK)	Introduced into Hangzhou Polytechnic's IoT program for perception/CV practicals; RoboUniversity certification courses	[deployed] (vocational-college/university classroom practicals)	[127138]
Jueying Lite3	DEEP Robotics	Education/research version roughly US$2,890 (one of the lowest-threshold research-grade platforms at present)	University/research teaching platform	[deployed] (education/research)	[130254]
Jueying X30 (industrial flagship)	DEEP Robotics	—	Industrial inspection (Wenzhou power-grid tunnels, Singapore SP Group); not education	[deployed] (industrial inspection)	[131]
BabyAlpha A2	Weilan	Base version RMB 8,099 → Ultra RMB 26,099 (built-in LLM learning/companionship)	A home consumer product with learning and entertainment, not a classroom teaching	[mass-produced] (consumer/home)	[133]

			product		
Spot	Boston Dynamics (US)	Explorer Kit from roughly US$74,500	As of 2025 Q4, 480+ universities/research institutions worldwide hold an academic license; industrial inspection	[deployed] (industry + university research)	[134]

> *Quadruped vs. humanoid education landing: quadruped locomotion control is engineered, the price threshold is low (Go2 EDU roughly US$5,990, Lite3 roughly US$2,890), and it* ***has already entered vocational-college/university classrooms for perception and algorithm practicals****; humanoids are mainly university research-and-teaching platforms, while at-scale primary/secondary use still relies on small programming robots.* ***Usage discipline****: one may write "the quadruped is the most realistic platform for embodied intelligence to enter education and has done perception and algorithm practicals in vocational colleges/universities," but one* ***must not write*** *"robot dogs have already entered primary and secondary classrooms at scale." The global quadruped market in 2025 is roughly US$541 million / roughly 18,500 units (a firm forecast caliber, to be annotated, not an education TAM).*

E.7 Cautious Conclusions from the Matrix Overview

Across the five model classes, three points can be distilled: (1) what is truly at scale in classrooms is still **programming/STEM kits and small social/teaching robots** (mature tracks of the pre-foundation-model era); (2) among the embodied frontier (humanoids/dexterous hands/quadrupeds), **only the quadruped has achieved classroom-practical landing in vocational colleges/universities**, while humanoids and dexterous hands remain mainly university research platforms; and (3) Western humanoids are industry-only, with no education deployment. Whole-machine specifications are vendor figures, not independently verified, and price ranges are not shelf quotes; the maturity tags and currency calibers must be retained when citing.

Cross-references: → Chapter 6 (body hardware), → Chapter 9 (deployment cases), → Chapter 13 (vendor financials/shipments), → Appendix F (the PENDING register of single-source financing/calibers). Whole-machine specifications are marked "vendor figures, not independently verified"; price ranges are not shelf quotes.

Appendix F Data Availability Statement and Open-Items List

> *This appendix discloses, in the form of a Nature-style Data Availability Statement (DAS), the source, availability, and caliber risk of all data in this blue book, and faithfully registers every unresolved item (PENDING). This blue book holds to a principle of transparent disclosure — it neither evades nor conceals any single-source, unverified, or caliber-conflicting item, so that readers may review it with care.*

F.1 Data Availability Statement (DAS)

All data supporting this blue book's conclusions come from **publicly available sources**; the study uses no restricted or confidential data and produces no original experimental data. Specifically, the data draw on public market-research-firm reports, original policy texts issued by governments and regulators, listed-company annual reports and prospectuses, peer-reviewed literature, and vendors' public specifications and official announcements. The complete reference list — **283 entries**, each carrying title, author or institution, year, source type, and an accessible link — is provided in the bibliography; every body-text citation is traceable to that library.

Data governance follows three principles:

- **Multi-source cross-verification**: all key market sizes, shipment figures, and financial numbers are cross-verified against no fewer than two mutually independent sources; any figure backed by fewer than two sources is explicitly marked "single source to be checked" in the body text and registered in the PENDING list in F.2 below. - **Side-by-side calibers and conservative adoption**: figures with caliber conflicts (e.g., broad vs. narrow caliber, divergent market forecasts from different firms) are presented side by side with the divergence explained, taking a conservative range — never mixing calibers or treating the most aggressive single value as settled. - **Currency-caliber checking**: wherever a Chinese company's financial figure is involved, the "亿元 ↔ billion" conversion follows the value checked against the original text, so as to avoid systematic misreading (see F.3).

Policy clauses follow the official original texts and official publication links (see Appendices C and D); market size, shipments, and financial figures come from third-party research-firm reports, listed-company annual reports and prospectuses, or government public documents. This blue book holds to transparent disclosure — every single-source, unverified, or caliber-conflicting item is faithfully registered in the PENDING list below, neither evaded nor concealed.

F.2 Open-Items List (Data Gaps and Items Pending Verification)

The following faithfully registers all unresolved items in four categories — data gaps, caliber conflicts, bibliographic fields to be supplemented, and source-credibility notes — for readers to review with care.

A. Data Gaps (No Cross-Source / Single Source to Be Checked)

#	Item	Status	Handling
A1	No cross-source for the independent USD size of sub-tracks: no credible cross-source was found for the 2026 independent market size of the social/companion, language-learning, and early-childhood sub-tracks	Gap	No figure force-assigned; the body text presents these qualitatively as "model deployment + academic evidence"
A2	Single-source financing/funding amounts awaiting a second check (PENDING-secondary): Makeblock's US$25 million Series D, VEX's US$18 million CTE funding, Fourier's financing of more than RMB 1 billion — all single media sources	Single source	Mark "single media source, amount awaiting a second check" when citing
A3	HolonIQ / iResearch / EqualOcean educational-robot dedicated size not hit: this round did not hit their direct dedicated caliber (most appear in the AI-education total rather than the educational-robot segment)	Gap	A targeted second-round search is recommended
A4	1X (NEO) education-path data thin: essentially judged "no clear education positioning"; its autonomous capability is actually remote teleoperation	Gap	To be supplemented
A5	Grand View report page body returns 403: the body	Restricted	The key figures have been cross-confirmed by three

	of the original page cannot be taken directly		parties — press release + PR Newswire + GII Research (conclusion reliable); citing the original table remains PENDING the original page
A6	Shanghai AI Fundamentals class-hour count (1 class-hour/week / at least 30/year): a secondary edtech-report caliber; "courses offered in grades 4/7" is an official brief (hard) but the class-hour count is not found in the original Shanghai Municipal Education Commission text	Single source	Recheck the original Shanghai Municipal Education Commission text before citing a hard class-hour count; mark the body text "per media reports"
A7	UBTech AI education covering "nearly 2,000 schools": a search-engine restatement of a vendor/annual-report caliber, without an audited breakdown; "5,000 schools" is a target value	Single source	Mark "vendor figures" when citing, and distinguish from the 5,000 target
A8	Whole-machine DoF/price/shipments mostly vendor figures: Tesla's mass-production price/capacity, various humanoids' shipments (prospectus vs. Omdia caliber differences)	Single source	Mark "vendor/media figures, not independently verified"

B. Caliber Conflicts (Presented Side by Side; Writers Must Annotate)

#	Item	Conflict	Handling
B1	2025 global humanoid shipments	Omdia 13,000 vs. IDC 18,000	Already side by side; if convergence is needed, Omdia is recommended as the main caliber and IDC as corroboration
B2	Global educational-robot CAGR	14.67% (Mordor) vs. 28.8% (GVR)	Take a range of 15%–29%, not the most aggressive

			28.8% as the sole caliber; the divergence turns on whether STEM kits/humanoids are included
B3	China broad caliber vs. narrow caliber	Broad roughly RMB 13.2 billion (132 亿元) vs. narrow roughly RMB 4.4 billion (44 亿元) (nearly threefold apart)	Never mix; both must be given and explained; Frost & Sullivan's 2030 narrow caliber of RMB 16.5 billion (165 亿元) can serve as a conservative anchor
B4	Product form / dominant school stage	Product: GVR says non-humanoid-dominated vs. Mordor says humanoids account for 49.13% (opposite conclusions); stage: GVR secondary vs. Mordor primary	Both must be marked as conflicting and presented side by side
B5	Dexterous-hand market-size caliber	2024 roughly US$815 million (Valuates, CAGR 40.4% aggressive single source) vs. 2025 roughly US$1.48 billion (another caliber)	Multi-firm calibers diverge widely; present side by side with care; not an education TAM; mark the 40% CAGR "aggressive single source"
B6	Quadruped 2025 shipments/market	Global roughly US$541 million / roughly 18,500 units (firm forecast); 2031 roughly 394,800 units (another caliber)	Mark "firm forecast, calibers diverge widely, not an education TAM"
B7	Unitree research-and-education share	Prospectus "roughly 74% sold to universities for research, roughly 9% into industry" vs. older table "73.6% from research and education" (2025.1–9)	Standardize the phrasing as "roughly three-quarters from research and education," with both calibers side by side

C. Bib Field Placeholders / To Be Supplemented

#	bib key	Issue	Handling
C1	`zhong2025asd_three_level`	The author field is "Zhong, and others"; volume and issue to be supplemented	Usable for citation; standardize the fields before formal publication
C2	`learnlm2025deepmind`	A DeepMind report;	Usable for citation;

		institutional-author placeholder; formal citation format to be standardized	standardize the format before formal publication
C3	Certain technical entries (RT-1/RT-2/OpenVLA/Octo/Gemini Robotics, etc.)	Share the VLA-survey URL as a citation source (not duplicate entries)	Tracing back to each original paper's DOI/arXiv is recommended at formal publication

D. Source-Credibility Notes

#	Item	Note
D1	Some historical URLs on the Global Educational Robots White Paper original-report site now return 404	Owing to a site redesign / anti-scraping; but the definitions, roles, seven-layer chain, and market figures have all been cross-verified against independent third-party reproductions (199IT/ScienceNet/Sina/Zhidx/36Kr/Jiemian, etc.) and match the local source files verbatim — no framework element was found that could not be traced
D2	Humanoid hardware whole-machine specifications (DoF/price/shipments)	Come mostly from vendor spec sheets and third-party aggregator sites, and must be marked "vendor figures, not independently verified," tracing back to vendor official sites/financials where possible
D3	MarketResearch.Biz cited via a Qianzhan Industry Research Institute Chinese restatement	The original firm was not directly verified; use with care
D4	EU Digital Omnibus legislative status	As of 2026-06, still not finally adopted: provisional political agreement 2026-05-07 + European Parliament endorsement 2026-06-16, with the formal text / OJ publication expected 2026-07; writers must mark "a proposal/provisional agreement, not finally in force; the baseline remains Reg (EU) 2024/1689"; the Art. 5(1)(f) emotion-recognition ban is already in force and unaffected
D5	Chinese anthropomorphic-AI measures' enforcement status	Promulgated 2026-04-10, effective 2026-07-15; as of 2026-06 not yet in force, no enforcement case, and "promulgated / to take effect" must be distinguished

F.3 Special Note on Currency Caliber (Highest Priority)

This study found a systematic risk that overseas search engines misread the Chinese "亿元" as "billion." Wherever a Chinese company's financial figure is involved, the "亿元 ↔ billion" conversion follows the value checked against the original (1 亿元 ≈ US$14 million, about 0.0014 billion — not 1 billion). The key errata resolved (including v2-refreshed values): UBTech's 2025 revenue **RMB 2.001 billion (20.01 亿元)** (+53.3%), UBTech's AI-education business **RMB 413 million (4.13 亿元)** (+13.7%, **replacing the old 2024 value of RMB 363 million (3.63 亿元)**), UBTech's 2025 full-size humanoids **RMB 821 million (8.21 亿元)**, Unitree's 2025 revenue **RMB 1.708 billion (17.08 亿元)** (not 17 billion), and iFLYTEK's 2025 smart education **RMB 8.967 billion (89.67 亿元)** (+24.04%, not 89.67 billion). Each value above was checked against the original vendor annual report or prospectus and cross-confirmed by no fewer than two independent sources.

F.4 Note on Historical-Forecast Calibers

The 2019 edition forecast "US$84.1 billion globally in 2023" and the 2016 edition forecast "US$11.1 billion globally in 2021"; both were research-team forecasts with too broad a caliber (mixing in consumer toys and training services) and **clearly overestimated** the market. The 2026 measurement: even the most aggressive GVR caliber reaches only US$5.84 billion in 2030. This blue book takes an honest look back at this, **strictly forbids reusing the US$84.1 billion / US$11.1 billion figures as accomplished facts**, and on that basis proposes the methodological improvement that "the narrow-scope body market and the broad STEAM consumer market must be counted separately."

Cross-references: → Appendix A (methodology), → Chapter 8 (sub-tracks), → Chapter 12 (the market total). *All unresolved items registered in this appendix are transparently disclosed, not evaded.*

Summary of PENDING entries: Category A 8 + Category B 7 + Category C 3 + Category D 5 = **23 in total** (8 more than the 15 at the W1 stage, adding 8 v2-upgrade unresolved items: Shanghai class-hours, nearly 2,000 schools, Fourier financing, whole-machine vendor figures, dexterous-hand/quadruped market calibers, Unitree share caliber, EU Omnibus legislative status, and the anthropomorphic-AI measures' enforcement status).

Appendix G References and Data Sources

> *The references and data sources of this blue book are gathered in a single bibliography, `01_research/references.bib` (283 BibTeX entries, expanded from the 133 entries of the W1 stage through v2 deep research). This appendix explains the citation system, the source-credibility tiering, and the composition of the bibliography; the "CITE placeholders" in the body text (in the double-brace form `CITE:key`) are resolved by the assembly script into numbered superscript endnotes, and the corresponding full reference entries are generated automatically after this appendix.*

G.1 The Citation System: CITE Placeholders and Assembly Resolution

This blue book uses a two-stage "placeholder + assembly" citation mechanism. At the drafting stage, the body text marks citations with CITE placeholders wrapped in double braces (in the form `CITE:key`), where the "key" must be a key already present in `references.bib`. At the assembly stage, the script `05_qa/assemble_robot_cn.py` matches each CITE placeholder with the regular expression `\{\{CITE:([\w\-]+)\}\}`, resolves it into that reference's citation number in the full text, presents it as a **superscript endnote**, and generates the complete reference list by number at the end of the document. Multiple citations of the same reference share a single number. This mechanism guarantees that: (1) citations are strongly consistent with the bibliography (a key must indeed exist in the library); (2) writers need not maintain numbers by hand; and (3) citation numbers in the whole volume are unique and continuous after assembly.

During the drafting stage, the few citations for which no entry yet existed were marked with a temporary comment rather than a fabricated bib key; by the final stage all had been entered as real, verifiable entries (such as the WEF Education 4.0 agenda in Appendix C and the 2017 State Council Next Generation Artificial Intelligence Development Plan in Appendix D). Placeholder entries whose bibliographic fields still require a second check (such as the three-level ASD meta-analysis and the LearnLM report) are registered together in Appendix F for readers to cite with care.

G.2 Composition of the Bibliography (283 Entries)

After merging and de-duplication, `references.bib` contains **283** entries (133 from the W1 stage + 150 added through v2 deep research), distributed by BibTeX entry type as follows:

Entry type	Count	Principal content
`@article`	45	Academic journal and conference papers: evidence-based meta-analyses and systematic reviews (STEM/CT, language learning,

		special education, early childhood), technical reviews (VLA, embodied intelligence, world models), and social-robot education research
`@misc`	232	Official policy text and documents, market-research-firm reports, listed-company annual reports and prospectuses, trade-press coverage, and vendor spec sheets and secondary aggregator sites (with v2-added entries on humanoids/dexterous hands/quadrupeds/world models/educational landing)
`@techreport`	4	The two historical editions of the Global Educational Robots White Paper (the historical-framework foundation) + reports such as DeepMind LearnLM
`@inproceedings`	2	Conference papers such as OpenVLA
Total	283	—

By subject matter, the bibliography covers the five research tracks and the three v2 deep-research tracks: **market and industry** (firm market reports, vendor financials/prospectuses, the financing landscape), **technology evolution** (papers and reviews on VLA/embodied intelligence/world models), **policy and governance** (the official policy text of 12 jurisdictions + international-organization frameworks + local embodied-industry policies), **evidence** (meta-analyses/systematic reviews across settings), and the three v2-added groups — **humanoid and dexterous-hand hardware/upstream components**, **quadruped education landing and world models**, and **real-world educational-landing cases**; the historical foundation is the two historical editions of the *Global Educational Robots White Paper* and the third-party-reproduction sources used for cross-verification.

G.3 Source-Credibility Tiering

To help readers judge the strength of the evidence, this blue book applies a four-tier credibility ranking to its sources, paired with the maturity tags in the body text and the annotations of the verification report:

1. **Peer-reviewed academic literature (highest)**: journal and conference papers, meta-analyses, and systematic reviews (such as `zhang2021robots_ct`, `plos2022asd_robots`, `vandenberghe2019language`), used for evidence-based conclusions, with effect sizes never fabricated and adverse effects recorded equally. 2. **Official policy text / government documents (high)**: laws, regulations, departmental rules, and government gazettes (such as `cac2026anthropomorphic`, `euaiact2024`, `moe2026aiedu`), with clauses and effective dates following the original. 3. **Firm market reports / listed-company annual**

reports and prospectuses (medium-high, requiring scope annotation): reports from Mordor, GVR, Research and Markets, MarketsandMarkets, and others, together with UBTech's annual report and Unitree's prospectus (such as `mordor-edurobot-2031`, `gvr-edurobot-market-2030`, `21jingji-ubtech-2024-revenue`, `eeo-unitree-prospectus-2026`); at least two cross-sources are required, conflicts are presented side by side, and currency calibers are checked back against the original. 4. **Trade press / secondary aggregator sites (requiring caution, flagged as not independently verified)**: vendor spec-sheet aggregations and media reproductions (such as `humanoid_compare2026`, `robozaps-humanoid-education-2026`, `qianzhan-global-edurobot-58usd-2025`); flagged "vendor figures / single media source, not independently verified," with single-source amounts entered into the PENDING list (see Appendix F).

G.4 Note on Placeholder and To-Be-Standardized Entries

In keeping with the principle of transparent disclosure, a small number of entries in the bibliography have field placeholders or citation sources to be standardized, and should be handled before formal publication (registered in parallel in Appendix F, Category C):

- `zhong2025asd_three_level`: the author field is "Zhong, and others"; volume and issue to be supplemented. - `learnlm2025deepmind`: a DeepMind report; institutional-author placeholder; formal citation format to be standardized. - Certain technical entries (RT-1/RT-2/OpenVLA/Octo/Gemini Robotics, etc.) share the VLA-survey URL as a citation source (not duplicate entries); tracing back to each original paper's DOI/arXiv is recommended at formal publication.

G.5 The Full Reference List

The complete list of 283 references is generated automatically from `references.bib` by the assembly script, numbered in order of first citation in the full text and corresponding one-to-one with the superscript endnotes in the body. Readers can use the numbers to trace each reference's title, author/institution, year, and URL.

Cross-references: → all chapters (citation sources), → Appendix A (methodology and verification), → Appendix F (PENDING and credibility notes). Placeholder entries must be standardized before formal publication.

References and Data Sources (in order of first appearance)

[1] 刘德建 and 黄荣怀 and 陈年兴 and 樊磊 and 赵国仁 and 沈锰坤 and 孙培真 and 周京明 and 李庆长 and 郑怡玲 and 郑雅文 and 邹诗薇 and 许荣宏 (2016). 全球教育机器人发展白皮书 (Educational Robots White Paper 2016: The Global Development). [in Chinese] https://sli.bnu.edu.cn/

[2] 北京师范大学 (2019). 2019 全球教育机器人发展白皮书 (Educational Robots White Paper 2019: The Global Development). [in Chinese] https://sli.bnu.edu.cn/en/News/SLI_News/2019/0830/741.html

[3] 199IT 中文互联网数据研究资讯中心 (2017). 北京师范大学: 2016 全球教育机器人发展白皮书（全文转载）. [in Chinese] https://www.199it.com/archives/640276.html

[4] Brohan, Anthony and others (2023). RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. https://en.wikipedia.org/wiki/Vision-language-action_model

[5] Kim, Moo Jin and others (2024). OpenVLA: An Open-Source Vision-Language-Action Model. https://en.wikipedia.org/wiki/Vision-language-action_model

[6] Black, Kevin and others (2024). π_0. https://arxiv.org/pdf/2410.24164

[7] NVIDIA (2025). GR00T N1: An Open Foundation Model for Generalist Humanoid Robots. https://arxiv.org/abs/2503.14734

[8] Figure AI (2025). Helix: A Vision-Language-Action Model for Humanoid Robotics. https://www.therobotreport.com/figure-humanoid-robots-demonstrate-helix-model-household-chores/

[9] 黄荣怀 and 陈莺 and Tlili, Ahmed (2024). 教育机器人的典型应用场景及技术治理. [in Chinese] https://aiedchair.bnu.edu.cn/

[10] Khan Academy (2024). How Khan Academy Is Building a Better AI Tutor (Khanmigo). https://blog.khanacademy.org/how-khan-academy-is-building-a-better-ai-tutor-our-most-recent-learnings/

[11] Khan Academy / Global Society (2025). Khan Academy rolls out AI-powered teaching tools as school districts scale up adoption.

[12] 国家网信办、国家发展改革委、工业和信息化部、公安部、市场监管总局 (2026). 人工智能拟人化互动服务管理暂行办法. [in Chinese] https://www.cac.gov.cn/2026-04/10/c_1777558395078289.htm (accessed 2026-06-25)

[13] European Commission AI Act Service Desk (2025). AI Act Article 5: Prohibited AI Practices. https://ai-act-service-desk.ec.europa.eu/en/ai-act/article-5 (accessed 2026-06-25)

[14] Federal Trade Commission (2025). FTC COPPA Rule — 2025 Amendments (biometric identifiers, opt-in). https://www.ftc.gov/legal-library/browse/rules/childrens-online-privacy-protection-rule-coppa (accessed 2026-06-25)

[15] Grand View Research (2026). Educational Robot Market Size, Share \& Trends Analysis Report, 2030.

[16] Research and Markets (2026). Educational Robot Market Report 2026-2030.

[17] 国务院 (2025). 国务院关于深入实施“人工智能+”行动的意见（国发〔2025〕11 号）. [in Chinese] https://www.news.cn/politics/20250826/21f5785636b14373af2e5d85ef383344/c.html (accessed 2026-06-25)

[18] 中共中央、国务院 (2025). 教育强国建设规划纲要（2024—2035 年）. [in Chinese] https://www.gov.cn/zhengce/202501/content_6999913.htm (accessed 2026-06-25)

[19] 教育部、国家发展改革委、工业和信息化部、科技部、国家数据局 (2026). “人工智能+教育”行动计划（教科信〔2026〕1 号）. [in Chinese] https://www.eol.cn/zhengce/wenjian/202604/t20260410_2727386.shtml (accessed 2026-06-25)

[20] 中华人民共和国教育部 (2026). 介绍《"人工智能+教育"行动计划》有关情况 - 新闻发布会. [in Chinese]

[21] 教育部基础教育教学指导委员会 (2025). 中小学人工智能通识教育指南（2025 年版）. [in Chinese] https://www.edu.cn/xxh/focus/zc/202505/t20250513_2667990.shtml (accessed 2026-06-25)

[22] 北京市人民政府（首都之窗） (2025). 中小学秋季学期起开设 AI. [in Chinese]

[23] 中国教育和科研计算机网（教育部） (2025). 教育部简报：上海中小学四年级、七年级开设《人工智能基础》地方课程. [in Chinese]

[24] Press Information Bureau / CBSE / NCERT (2025). India CBSE AI \& Computational Thinking Curriculum (Class 3-12). https://www.pib.gov.in/PressReleasePage.aspx?PRID=2184211 (accessed 2026-06-25)

[25] Singapore Ministry of Education (2025). Singapore MOE — Artificial Intelligence in Education (EdTech Masterplan 2030). https://www.moe.gov.sg/education-in-sg/educational-technology-journey/edtech-masterplan/artificial-intelligence-in-education (accessed 2026-06-25)

[26] UK Department for Education (2025). DfE Generative AI in schools and colleges guidance. https://educationhub.blog.gov.uk/2025/06/artificial-intelligence-in-schools-everything-you-need-to-know/ (accessed 2026-06-25)

[27] Japan MEXT (2025). Guideline for the Use of Generative AI in Primary and Secondary Education. https://www.mext.go.jp/content/20250422-mxt_shuukyo01-000030823_001.pdf (accessed 2026-06-25)

[28] Australian Government Department of Education (2024). Australian Framework for Generative AI in Schools. https://www.education.gov.au/schooling/resources/australian-framework-generative-artificial-intelligence-ai-schools (accessed 2026-06-25)

[29] 工业和信息化部 (2023). 工业和信息化部《人形机器人创新发展指导意见》（工信部科〔2023〕193 号）. [in Chinese]

[30] 深圳市科技创新局 (2025). 《深圳市具身智能机器人技术创新与产业发展行动计划（2025—2027 年）》. [in Chinese]

[31] 上海市人民政府办公厅 (2025). 《上海市具身智能产业发展实施方案》. [in Chinese]

[32] 北京市科委·中关村管委会 (2025). 《北京具身智能科技创新与产业培育行动计划（2025—2027 年）》. [in Chinese]

[33] 深圳新闻网 (2026). 具身智能赛道迎来上市潮 深圳人形机器人企业加速登陆资本市场. [in Chinese]

[34] 国家互联网信息办公室 (2026). 《人工智能拟人化互动服务管理暂行办法》答记者问. [in Chinese] http://www.hunan.gov.cn/zqt/zcjd/bmjd/202604/t20260413_33953005.html (accessed 2026-06-25)

[35] 教育部基础教育教学指导委员会 (2025). 中小学生成式人工智能使用指南（2025 年版）. [in Chinese] https://www.canedu.org.cn/site/content/7876.html (accessed 2026-06-25)

[36] European Parliament and Council (2024). Regulation (EU) 2024/1689 (Artificial Intelligence Act) — Article 5 \& Annex III. https://artificialintelligenceact.eu/annex/3/ (accessed 2026-06-25)

[37] California Lawyers Association (overview) (2025). California SB 243 Companion Chatbots Act \& New York AI companion law. https://calawyers.org/privacy-law/regulatory-focus-on-ai-companion-character-chatbots/ (accessed 2026-06-25)

[38] New York State / Governor's Office (2025). New York AI Companion Models law (GBL §1700 et seq.) and S9051B.

[39] European Parliament and Council (Regulation (EU) 2016/679) (2018). GDPR Article 8 — Conditions applicable to child's consent. https://gdpr-info.eu/art-8-gdpr/ (accessed 2026-06-25)

[40] U.S. Department of Education (2024). U.S. Department of Education — Protecting Student Privacy (FERPA, AI guidance, NETP 2024). https://studentprivacy.ed.gov/ (accessed 2026-06-25)

[41] 国务院 (2023). 未成年人网络保护条例（国务院令第 766 号）. [in Chinese] https://www.gov.cn/zhengce/content/202310/content_6911288.htm (accessed 2026-06-25)

[42] 国家网信办等七部门 (2023). 生成式人工智能服务管理暂行办法. [in Chinese] https://www.gov.cn/zhengce/zhengceku/202307/content_6891752.htm (accessed 2026-06-25)

[43] Gibson Dunn / European Parliament Legislative Train (2026). EU Digital Omnibus on AI — Postponed High-Risk Deadlines. https://www.gibsondunn.com/eu-ai-act-omnibus-agreement-postponed-high-risk-deadlines-and-other-key-changes/ (accessed 2026-06-25)

[44] UNESCO (2024). UNESCO AI Competency Framework for Students. https://www.unesco.org/en/articles/ai-competency-framework-students (accessed 2026-06-25)

[45] OECD (2024). OECD AI Principles (2019, updated 2024) \& K-12 AI Literacy Framework. https://oecd.ai/en/ai-principles (accessed 2026-06-25)

[46] 智东西内参 (2017). 全球教育机器人深度报告，一文看尽 7 层产业链 12 类产品. [in Chinese] https://www.zhidx.com/p/164878.html

[47] 王建虎 and 王静 and 万亚玲 (2025). 具身智能赋能教育的内涵、框架与研究路向. [in Chinese] https://aver.nwnu.edu.cn/

[48] 黄荣怀 and 刘梦彧 and 刘嘉豪 and 张定文 (2023). 智慧教育之"为何"与"何为"——关于智能时代教育的表现性与建构性特征分析. [in Chinese]

[49] Survey (2025). Vision-Language-Action Models: Concepts, Progress, Applications and Challenges. https://arxiv.org/pdf/2505.04769

[50] Survey (2025). Vision Language Action Models in Robotic Manipulation: A Systematic Review. https://arxiv.org/pdf/2507.10672

[51] Authors (2017). Robotics as an Educational Tool: Impact of Lego Mindstorms. https://www.researchgate.net/publication/305876097_Robotics_as_an_Educational_Tool_Impact_of_Lego_Mindstorms

[52] IEEE Spectrum (2013). Nao Robot Goes to School to Help Kids With Autism (ASK NAO). https://spectrum.ieee.org/aldebaran-robotics-nao-robot-autism-solution-for-kids

[53] Authors (2024). Surprising Performances of Students with Autism in Classroom with NAO Robot. https://arxiv.org/html/2407.12014v1

[54] PYMNTS (2021). SoftBank Puts Pepper Robot On Hold (Discontinued 2021). https://en.wikipedia.org/wiki/Pepper_(robot)

[55] Physical Intelligence (2025). $\pi_0.5. https://www.pi.website/blog/pi05

[56] Brohan, Anthony and others (2022). RT-1: Robotics Transformer for Real-World Control at Scale. https://en.wikipedia.org/wiki/Vision-language-action_model

[57] Octo Model Team (2024). Octo: An Open-Source Generalist Robot Policy. https://en.wikipedia.org/wiki/Vision-language-action_model

[58] Google DeepMind (2025). Gemini Robotics. https://deepmind.google/models/gemini-robotics/

[59] Google DeepMind (2025). Gemini Robotics On-Device. https://deepmind.google/blog/gemini-robotics-on-device-brings-ai-to-local-robotic-devices/

[60] TechTimes (2026). The Data Drought: Why Embodied AI Can't Just Read the Internet.

[61] EVS International (2026). Embodied AI Data Collection: Teleoperation, Sim-to-Real Guide.

[62] Shaip (2025). Robot Training Data Strategy: Teleoperation vs Simulation vs Human Video for Embodied AI.

[63] World Labs (2026). Research and Insights — World Labs Blog (world model taxonomy essay, 2026-06-03).

[64] Google DeepMind (2025). Genie 3: A new frontier for world models.

[65] Google DeepMind (2026). Genie — Google DeepMind Model Page (Project Genie demo).

[66] Decart AI (2024). Oasis: A Universe in a Transformer.

[67] Decart AI (2025). MirageLSD: The First Live-Stream Diffusion AI Video Model.

[68] TechCrunch (2026). Decart's new world model can simulate hours of photorealistic driving — with some caveats.

[69] TechCrunch (2025). Odyssey's new AI model streams 3D interactive worlds.

[70] Odyssey (2025). Introducing Explorer / World Models for Film, Gaming, and Beyond.

[71] UlazAI (2025). AI Video Models Guide 2025: Runway GWM-1, Luma, Kling, Sora.

[72] TechCrunch (2025). Luma releases a new AI model that lets users generate a video from a start and end frame.

[73] NVIDIA Newsroom (2025). NVIDIA Announces Major Release of Cosmos World Foundation Models and Physical AI Data Tools.

[74] NVIDIA (2025). NVIDIA Cosmos: World Foundation Models Powering Physical AI.

[75] NVIDIA Technical Blog (2025). Scale Synthetic Data and Physical AI Reasoning with NVIDIA Cosmos World Foundation Models.

[76] Wayve (2025). GAIA-2: Pushing the Boundaries of Video Generative Models for Safer Assisted and Automated Driving.

[77] Wayve (2025). GAIA-2: A Controllable Multi-View Generative World Model for Autonomous Driving.

[78] Meta AI (2025). Introducing the V-JEPA 2 world model and new benchmarks for physical reasoning.

[79] Meta AI (2025). V-JEPA 2: Self-Supervised Video Models Enable Understanding, Prediction and Planning.

[80] TechCrunch (2025). Fei-Fei Li's World Labs speeds up the world model race with Marble, its first commercial product.

[81] TechTimes (2026). Embodied AI World Models Attracted \$6 Billion, But the LLM Parallel May Not Hold.

[82] arXiv (2025). A Comprehensive Survey on World Models for Embodied AI.

[83] NVIDIA Blog (2025). Into the Omniverse: Open World Foundation Models Generate Synthetic Worlds for Physical AI Development.

[84] RoboStore (2026). Unitree G1 (U2) EDU Plus Robotic Humanoid.

[85] New Hampshire Department of Education (2025). Khan Academy to extend its AI services, at no cost, to New Hampshire educators and students.

[86] Iowa Department of Education (2024). Iowa Department of Education launches new personalized reading tutor (Amira).

[87] PR Newswire / Amira Learning (2026). Ministry of Education of Jordan begins national literacy initiative using Amira Learning.

[88] Google (2025). Google for Education year in review 2025.

[89] Authors (2024). Enhancing Educational Dynamics: Integrating Large Language Models with a Social Robot. https://dl.acm.org/doi/full/10.1145/3715885.3715889

[90] Authors (2025). The Potential of Large Language Models for Social Robots in Special Education. https://link.springer.com/article/10.1007/s13748-025-00363-2

[91] Authors (2025). A Humanoid Social Robot as a Teaching Assistant in the Classroom. https://arxiv.org/html/2508.05646v1

[92] Authors (2025). RoboBuddy in the Classroom: Exploring LLM-Powered Social Robots for Storytelling. https://www.researchgate.net/publication/394940844_RoboBuddy_in_the_Classroom

[93] Authors (2024). Fostering Children's Creativity through LLM-Driven Storytelling with a Social Robot. https://pmc.ncbi.nlm.nih.gov/articles/PMC11671368/

[94] Authors (2025). Robots and Children that Learn Together: Improving Knowledge Retention by Teaching Peer-Like Interactive Robots. https://arxiv.org/pdf/2506.18365

[95] LearnLM Team and Eedi (2025). AI tutoring can safely and effectively support students: An exploratory RCT in UK classrooms.

[96] LumiChats (2026). Humanoid Robots 2026: Tesla Optimus vs Figure AI vs Unitree. https://lumichats.com/blog/humanoid-robots-2026-tesla-optimus-figure-ai-unitree-complete-guide

[97] AI Business (2026). What a Humanoid Robot Actually Costs in 2026 (Unitree vs Optimus vs Figure). https://aibusiness.vc/robots/humanoid-robot-price-comparison-2026

[98] (2026). Tesla Optimus Gen 3: Specs, Production Timeline and Cost Targets. https://airobots.media/technology/tesla-optimus-gen-3-everything-we-know-about-teslas-most-ambitious-product/

[99] Figure AI (2025). Figure 02/03 and Helix at BMW Spartanburg; BotQ Mass Production. https://www.figure.ai/news/production-at-bmw

[100] 1X Technologies (2025). 1X NEO Home Humanoid: Preorder, 22-DoF Hands, Teleoperation Controversy. https://www.therobotreport.com/1x-announces-pre-order-launch-neo-humanoid-robot/

[101] (2025). Teleop, Not Autonomy, Is the Path for 1X's NEO Humanoid. https://www.therobotreport.com/teleop-not-autonomy-the-path-for-1x-neo-humanoid/

[102] Apptronik (2025). Apptronik Apollo: Specs and Mercedes-Benz/GXO/Jabil Deployments. https://apptronik.com/apollo

[103] Boston Dynamics (2026). Boston Dynamics Electric Atlas: 56 DoF Production Version, Hyundai/DeepMind Commitments. https://bostondynamics.com/blog/boston-dynamics-unveils-new-atlas-robot-to-revolutionize-industry/

[104] Agility Robotics (2026). Agility Robotics Digit v4/v5: Commercial Warehouse Deployment. https://www.agilityrobotics.com/content/digit-moves-over-100k-totes

[105] Sanctuary AI (2025). Sanctuary AI Phoenix: 21-DoF Hydraulic Hand with Tactile Sensors. https://www.therobotreport.com/sanctuary-ai-integrates-tactile-sensors-into-phoenix-general-purpose-robots/

[106] Unitree Robotics (2025). Unitree G1/H1/H2/R1 Humanoid Specs and Shipments. https://www.unitree.com/cn/mobile/news/28/

[107] 经济观察网 (2026). 深度拆解宇树科技招股书：一年营收 17 亿，宇树机器人都卖给了谁？. [in Chinese]

[108] 智元创新（上海） (2025). AgiBot (Zhiyuan) Yuanzheng A2 and Lingxi X2 Specs and 5000th Unit. [in Chinese] https://www.zhiyuan-robot.com/products/X2

[109] Fourier Intelligence (2025). Fourier GR-2 and GR-3 Care-Bot Humanoid Specs. https://botinfo.ai/articles/fourier-humanoid-robots

[110] UBTECH Robotics (2025). UBTECH Walker S2: Autonomous Battery Swap, Mass Production and Orders. https://www.prnewswire.com/news-releases/ubtech-humanoid-robot-walker-s2-begins-mass-production-and-delivery-with-orders-exceeding-800-million-yuan-302616924.html

[111] 北京星动纪元 (2025). Robot Era (星动纪元) STAR1 Humanoid Specs and Financing. [in Chinese] https://www.robotera.com/

[112] EngineAI Robotics (2025). EngineAI PM01 and SE01 Humanoid for Commercial and Educational Use. https://www.therobotreport.com/engineai-releases-pm01-humanoid-robot-for-commercial-educational-use/

[113] Booster Robotics (2025). Booster Robotics T1/K1: RoboCup Champion Education/Research Humanoid. https://www.booster.tech/booster-t1/

[114] 北京银河通用机器人 (2025). Galbot G1 Wheeled Humanoid: Retail/Pharmacy Deployment and Peking Univ. Lab. [in Chinese] https://www.galbot.com/g1

[115] (2025). 松延动力 N2/E1 教育科研人形机器人订单. [in Chinese] https://news.caijingmobile.com/article/detail/554046

[116] Shadow Robot Company (2025). Shadow Dexterous Hand Technical Specification. https://shadowrobot.com/wp-content/uploads/2025/09/shadow_dexterous_hand_lite_technical_specification-2025.pdf

[117] Wonik Robotics (2025). Wonik Robotics Allegro Hand V4/V5 Specs. https://www.allegrohand.com/

[118] Sharpa Robotics (2025). Sharpa Wave: 22-DoF Dexterous Hand with Dynamic Tactile Array, Mass Production. https://www.prnewswire.com/news-releases/ai-robotmaker-sharpa-reaches-key-milestone-with-mass-production-of-worlds-most-advanced-human-sized-robotic-hand-302643434.html

[119] 北京因时机器人 (2025). 因时机器人 Inspire-Robots RH56/RH5EG1 五指灵巧手. [in Chinese] https://www.inspire-robots.com/dexterous%20hands/rh56dfx-series/

[120] BrainCo 强脑科技 (2025). BrainCo Revo2 Bionic Dexterous Hand. [in Chinese] https://www.prnewswire.com/news-releases/brainco-launches-revo2-hand-for-humanoid-robotics-302557714.html

[121] 灵心巧手 LinkerBot (2025). LinkerBot Linker Hand L20/L25 Multi-Sensor Dexterous Hand. [in Chinese] https://www.linkerbot.cn/

[122] PaXini 帕西尼 (2026). PaXini DexH13 Dexterous Hand with Multidimensional Tactile Infrastructure. [in Chinese] https://www.prnewswire.com/news-releases/paxini-unveils-the-tactile-infrastructure-for-embodied-ai-redefining-full-stack-product-matrix-at-ces-2026-302655238.html

[123] OYMotion 傲意科技 (2025). OYMotion ROHand (ROH-A002/AP001/LiteS001) Dexterous Hand. [in Chinese] https://www.oymotion.com/product61

[124] 甲子光年 (2025). IROS 灵巧手大盘点：特斯拉还没做到的，中国厂商做到了. [in Chinese] https://www.jazzyear.com/article_info.html?id=1600

[125] (2025). 人形机器人行星滚柱丝杠国产化与成本占比. [in Chinese] https://www.aibangbots.com/a/2259

[126] (2025). 2025 中国人形机器人六维力传感器市场分析. [in Chinese] https://www.cls.cn/detail/2131477

[127] RoboStore (2026). Unitree Go2 EDU AI Quadruped Robot Dog (40 TOPS) — Specs and Pricing.

[128] Airpuria (2026). Unitree GO2 Price Guide (Robot Dog).

[129] STEMfinity (2026). Unitree Go2 EDU AI Quadruped Robot Dog — Education Reseller.

[130] AwesomeRobots (2025). DEEP Robotics Lite3 — Quadruped Robot Specs and Review (from \$2,890).

[131] DEEP Robotics (2025). DEEP Robotics X30 — Official Product Page.

[132] TechNode (2025). Deep Robotics delivers X30 quadruped robot ``SPock" for Singapore Power inspections.

[133] Readhub (2024). 蔚蓝 BabyAlpha A2 机器狗发售：国内首发多模态交互能力，6099 元起. [in Chinese]

[134] Boston Dynamics (2025). Spot — Official Product Page; Academia and Education.

[135] InRobots (2025). Boston Dynamics Robot Dog Spot: Price, Features, and Real-World Uses.

[136] The Pricer (2025). How Much Does A Boston Dynamics Robot Dog Cost? (incl. ANYmal comparison).

[137] Ghost Robotics (2025). Vision 60 Q-UGV — Official Product Page.

[138] 杭州科技职业技术学院 (2025). 科技赋能教育 | 智能机器狗打造 AI 人才培养创新课堂（宇树 Go2 进物联网技术学院）. [in Chinese]

[139] 科大讯飞智慧教育 (2025). 和济科技节：虚拟数字人携四足机器狗炫酷登场未来课堂. [in Chinese]

[140] Grand View Research (2025). Educational Robots Market Size, Share — Industry Report.

[141] Grand Research Store (). Quadruped Robot Market Report, Size, Share 2026 forecast to 2034.

[142] 经济观察网 (2026). 深度拆解宇树科技招股书：一年营收 17 亿，宇树机器人都卖给了谁？. [in Chinese]

[143] Technerdo (2026). Humanoid Robots in 2026: Market Leaders, Deployments, and What Comes Next.

[144] Authors (2025). Integrating Emotional Intelligence, Memory Architecture, and Gestures for Empathetic Humanoid Robot Interaction in Education. https://www.frontiersin.org/journals/robotics-and-ai/articles/10.3389/frobt.2025.1635419/full

[145] Authors (2024). Exploring the Impact of Robot Interaction on Learning Engagement: A Comparative Study of Two Multi-modal Robots. https://slejournal.springeropen.com/articles/10.1186/s40561-024-00362-1

[146] Authors (2025). Student Engagement Assessment Using Multimodal Deep Learning. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0325377

[147] Authors (2025). MathBuddy: A Multimodal System for Affective Math Tutoring. https://arxiv.org/pdf/2508.19993

[148] Authors (2025). Generative AI-Powered Social Robots in Education: Opportunities and Challenges from a Delphi Study. https://www.tandfonline.com/doi/full/10.1080/0144929X.2025.2604060

[149] Ferrer, I. and Aymerich-Franch, L. (2026). Social robots for education: a global analysis of deployments in real scenarios.

[150] Vogt, Paul and others (2019). Second Language Tutoring Using Social Robots: A Large-Scale Study.

[151] Robozaps (2026). Humanoid Robots in Education 2026.

[152] Intel Market Research (2026). Robotics Education Market Outlook 2026-2034.

[153] Dataintelo (2026). STEM Robotics Kits Market Research Report 2034.

[154] Industry Research.biz (2026). STEM Robotics Kits Market Trends, Size \& CAGR of 5.5\%.

[155] Hong and others (2024). The impact of educational robots on students' computational thinking: A meta-analysis of K-12. https://doi.org/10.1007/s10639-023-12415-y (accessed 2026-06-25)

[156] (2026). Social robots for education: a global analysis of deployments in real scenarios. https://link.springer.com/article/10.1007/s11423-026-10611-4

[157] Randall, Natasha (2019). Robot-Assisted Language Learning: A systematic review. https://pmc.ncbi.nlm.nih.gov/articles/PMC9613956/ (accessed 2026-06-25)

[158] van den Berghe, Rianne and Verhagen, Josje and Oudgenoeg-Paz, Ora and van der Ven, Sanne and Leseman, Paul (2019). Social Robots for Language Learning: A Review. https://doi.org/10.3102/0034654318821286 (accessed 2026-06-25)

[159] 前瞻产业研究院 (2025). 预见 2025：2025 年中国教育机器人行业全景图谱. [in Chinese]

[160] Mordor Intelligence (2026). Educational Robot Market Size, Industry Growth, Trends \& Share Report 2031.

[161] PLOS One (2022). The use of social robots with children and young people on the autism spectrum: A systematic review and meta-analysis. https://doi.org/10.1371/journal.pone.0269800 (accessed 2026-06-25)

[162] (2022). The use of social robots with children and young people on the autism spectrum: A systematic review and meta-analysis. https://pmc.ncbi.nlm.nih.gov/articles/PMC9216612/

[163] 教育界网 (2025). 优必选 4.13 亿教育收入背后，藏着什么野心？. [in Chinese]

[164] 北京市人民政府 / 北京市教委 (2025). 《北京市推进中小学人工智能教育工作方案（2025—2027 年）》. [in Chinese]

[165] 腾讯新闻 / 优必选 (2024). 优必选科技：AI. [in Chinese]

[166] 新浪财经 (2026). 优必选 2025 年报点评：营收稳健增长，人形机器人产品迈向商业化应用. [in Chinese]

[167] 北京日报 (2025). 人形机器人``天工行者"爆火！4 个月订单破百. [in Chinese]

[168] Gasgoo (2026). UBTECH.

[169] InfoQ (2026). 宇树科技靠融资还是靠卖货活着？翻完招股书，我们挖到了这 8 个关键点. [in Chinese]

[170] 极客公园 (2026). 宇树首度披露招股书：2025 年净利润 6 亿，募资 42 亿重点投入``机器人大脑". [in Chinese]

[171] 中国教育新闻网（教育部） (2025). 人工智能赋能教育取得阶段性进展（央馆课程、育小苗智能体、超 1000 门精品课程）. [in Chinese]

[172] 中央电化教育馆 (2025). 教育部教育技术与资源发展中心（中央电化教育馆）首页通知（央馆人工智能课程/虚拟实验）. [in Chinese]

[173] 中华人民共和国教育部 (2026). 介绍《``人工智能+教育"行动计划》有关情况——新闻发布会（509 基点校、7 省/20 地市/18 高校）. [in Chinese]

[174] 中国高等教育学生信息网 (2026). 教育部等五部门关于印发《``人工智能+教育"行动计划》的通知（转载）. [in Chinese]

[175] 科学网 (2026). 本科专业上新！新增具身智能、脑机科学与技术等专业. [in Chinese]

[176] 网易 (2026). 2026 将新增 57 个专业，人工智能被重点提及（343 所机器人工程专业）. [in Chinese]

[177] 网易 (2025). 教育部公布全国 102 所高职院校成功备案智能机器人技术专业名单（2025 年招生）. [in Chinese]

[178] 量子位 (2020). 优必选科技联合杭州市余杭区落地人工智能教育项目，100 所中小学+2 大基地正式授课. [in Chinese]

[179] 中国职业技术教育网 (). 全国机械行业工业机器人与智能装备职教集团：一体化建设促进职业院校工业机器人技术专业发展. [in Chinese]

[180] K-12 Dive (2025). 3 questions for K-12 leaders to consider amid the AI tutoring boom.

[181] Khan Academy (2025). Khan Academy Annual Report SY24-25.

[182] North Dakota Department of Public Instruction (). Amira Learning (statewide K-5 literacy).

[183] Chalkbeat Newark (2025). Newark adopts AI literacy screener Amira Learning per state plan.

[184] The Robot Report (2025). Aldebaran, maker of Pepper and Nao robots, put in receivership.

[185] Association for Advancing Automation (A3) (2025). Shenzhen-based Maxvision secures rights to Pepper and Nao following Aldebaran auction.

[186] Robotics and Automation News (2025). RobotLAB pledges zero disruption to servicing Aldebaran robots despite liquidation.

[187] European Commission / CORDIS (2016). L2TOR — Second Language Tutoring using Social Robots (H2020 #688014).

[188] PR Newswire / Luxembourg Institute of Health (2025). LuxAI, Luxembourg Institute of Health and University of Birmingham launch first large-scale study (QTrobot, 69 families).

[189] Axios (2024). Maker of AI robots for kids abruptly shutters (Embodied/Moxie).

[190] OECD.AI Incidents (2024). Embodied Shutdown Bricks Moxie AI Robots.

[191] Van Robotics / authors (2025). ABii at School: Findings from a Long-Term In-School Field Study.

[192] 前瞻经济学人 (2025). 2025 年中国教育机器人产业供应链十大代表性企业. [in Chinese]

[193] Wang, Feifei and Cheung, Alan C. K. (2025). Robots' Social Behaviors for Language Learning: A Systematic Review and Meta-Analysis. https://doi.org/10.3102/00346543231216437 (accessed 2026-06-25)

[194] Zhang, Yanjun and Luo, Ronghua and Zhu, Yijin and Yin, Yuan (2021). Educational Robots Improve K-12 Students' Computational Thinking and STEM Attitudes: Systematic Review. https://doi.org/10.1177/0735633121994070 (accessed 2026-06-25)

[195] International Journal of STEM Education (2024). The effects of educational robotics in STEM education: a multilevel meta-analysis. https://doi.org/10.1186/s40594-024-00469-4 (accessed 2026-06-25)

[196] Benitti, Fabiane Barreto Vavassori (2012). Exploring the educational potential of robotics in schools: A systematic review. https://doi.org/10.1016/j.compedu.2011.10.006 (accessed 2026-06-25)

[197] PMC11259833 (2024). Enhancing computational thinking in early childhood education with educational robotics: A meta-analysis. https://pmc.ncbi.nlm.nih.gov/articles/PMC11259833/ (accessed 2026-06-25)

[198] Lee, Hansol and Lee, Jang Ho (2022). The effects of robot-assisted language learning: A meta-analysis. https://www.sciencedirect.com/science/article/abs/pii/S1747938X21000488 (accessed 2026-06-25)

[199] Educational Research Review (2025). Meta-analyzing the impacts of social robots for children's language development: Insights from two decades of research from 2003 to 2023. https://www.sciencedirect.com/science/article/abs/pii/S1747938X25000399 (accessed 2026-06-25)

[200] International Journal of Human-Computer Interaction (2025). Effects of Early Social Robot-Based Interventions for Social-Emotional Development in Children with Autism Spectrum Disorders and Typically Developing Children: A Meta-Analysis. https://doi.org/10.1080/10447318.2025.2587245 (accessed 2026-06-25)

[201] Zhong, and others (2025). The Role of Social Robot in Enhancing Social-Emotional Skill Development in Children With Autism: A Three-Level Meta-Analysis. https://doi.org/10.1111/jcal.70154 (accessed 2026-06-25)

[202] (2025). A Review of Socially Assistive Robotics in Supporting Children with Autism Spectrum Disorder. https://www.mdpi.com/2414-4088/9/9/98

[203] (2026). Utilizing Human-Robot Interaction in Autism Therapy to Enhance Children's Social Skills. https://dl.acm.org/doi/10.1145/3776539

[204] Leite, Iolanda and Martinho, Carlos and Paiva, Ana (2013). Social Robots for Long-Term Interaction: A Survey. https://doi.org/10.1007/s12369-013-0178-y (accessed 2026-06-25)

[205] International Journal of Social Robotics (2024). The Child Factor in Child-Robot Interaction: Discovering the Impact of Developmental Stage and Individual Characteristics. https://arxiv.org/html/2404.13432v1 (accessed 2026-06-25)

[206] Frontiers in Robotics and AI (2022). Social Robots in Applied Settings: A Long-Term Study on Adaptive Robotic Tutors in Higher Education. https://doi.org/10.3389/frobt.2022.831633 (accessed 2026-06-25)

[207] AI \& SOCIETY (2021). Can communication with social robots influence how children develop empathy? Best-evidence synthesis. https://doi.org/10.1007/s00146-021-01214-z (accessed 2026-06-25)

[208] Richter, Markus and others (2025). Leveraging Large Language Models for Robot-Assisted Learning of Morphological Structures in Preschool Children with Language Vulnerabilities. https://arxiv.org/abs/2509.22287 (accessed 2026-06-25)

[209] Wang, Rose E. and others (2024). Tutor CoPilot: A Human-AI Approach for Scaling Real-Time Expertise. https://arxiv.org/abs/2410.03017 (accessed 2026-06-25)

[210] Google DeepMind (2025). LearnLM: Improving Gemini for Learning (Nov 2025 report). https://storage.googleapis.com/deepmind-media/LearnLM/learnLM_nov25.pdf (accessed 2026-06-25)

[211] arXiv 2407.09975 (2024). The GPT Surprise: Offering Large Language Model Chat in a Massive Coding Class Reduced Engagement but Increased Adopters' Exam Performances. https://arxiv.org/abs/2407.09975 (accessed 2026-06-25)

[212] 水清木华研究中心 (2026). 2026 年具身智能机器人大模型（含 VLA）研究报告. [in Chinese]

[213] AI Insight (2026). 具身智能产业全景 2026：谁在赚钱，谁在烧钱，谁在量产. [in Chinese]

[214] (2025). 2025 谐波减速器行业分析：人形机器人助力国产化率突破 30%. [in Chinese] https://www.baogaobox.com/insights/250518000010183.html

[215] (2025). 人形机器人无框力矩电机市场格局. [in Chinese] https://www.aibangbots.com/a/1764

[216] (2025). 国产电子皮肤/触觉传感器产业链全景. [in Chinese] https://www.aibangbots.com/a/4289

[217] Grand View Research (2026). Educational Robot Market Size To Reach \$5.84Bn By 2030.

[218] PR Newswire / Grand View Research (2023). Educational Robot Market to Hit \$5.5 Billion by 2030: Grand View Research, Inc..

[219] MarketsandMarkets (2022). Educational Robot Market - Global Forecast.

[220] MarketsandMarkets / PR Newswire (2021). Educational Robot Market worth \$2.6 billion by 2026 - Exclusive Report by MarketsandMarkets.

[221] 百谏方略 (DIResearch) (2025). 中国教育机器人专精特新企业细分市场占有率专项调研报告. [in Chinese]

[222] 前瞻产业研究院 (2025). 2025 年全球教育机器人市场分析 2030 年市场规模将上升至 58 亿美元. [in Chinese]

[223] 智研咨询 (2025). 2025 年中国教育机器人行业发展现状、重点企业经营情况及发展趋势研判. [in Chinese]

[224] 华经产业研究院 (2025). 2025 年中国教育机器人行业市场规模、相关政策、产业链及发展趋势分析. [in Chinese]

[225] Bloomberg / Omdia (2026). Chinese Firms Dominated Global Humanoid Robot Shipments in 2025.

[226] TechNode / Omdia (2026). China's AgiBot leads global humanoid robot shipments in 2025, Omdia says.

[227] CGTN / IDC (2026). IDC report: China leads the global humanoid robot rise in 2025.

[228] CNBC (2026). Morgan Stanley doubles China humanoid robot shipment forecast as commercialization accelerates.

[229] Goldman Sachs (). The global market for humanoid robots could reach \$38 billion by 2035.

[230] Goldman Sachs (2024). The global market for humanoid robots could reach \$38 billion by 2035.

[231] Morgan Stanley (2025). Humanoid Robot Market Expected to Reach \$5 Trillion by 2050.

[232] Valuates Reports / PR Newswire (). Dexterous Hands Market Size to Reach USD 10.3 Billion by 2031, Driven by Humanoid and Industrial Robots.

[233] 360iResearch (). Robot Multi-fingered Dexterous Hand Market 2025-2030.

[234] 科学网 (2016). 全球首部教育机器人白皮书正式发布——未来 5 年全球市场规模或达百亿美元. [in Chinese] https://news.sciencenet.cn/sbhtmlnews/2016/9/316094.shtm

[235] 新浪教育 (2019). 《2019 全球教育机器人发展白皮书》发布 提出 6 大核心观点. [in Chinese] http://edu.sina.com.cn/l/2019-08-23/doc-ihytcitn1327133.shtml

[236] 腾讯新闻 (2026). 人形销量破千，优必选如何成就全球第一. [in Chinese]

[237] 新浪财经 / 慧聪教育 (2025). 科大讯飞 2025 上半年营收破百亿，智慧教育业务贡献超三成. [in Chinese]

[238] 界面新闻 (). 教育业务已成为科大讯飞第一大业务，``十四五"末营收目标 300 亿. [in Chinese]

[239] 前瞻产业研究院 (2025). 启示 2025：中国教育机器人行业投融资及兼并重组分析. [in Chinese]

[240] 智教新媒 (2026). 万字拆解《2026 具身智能与人形机器人产业研究报告》：324 起融资、390 亿热钱涌入. [in Chinese]

[241] 财联社 (2025). 规模化量产时代来临？智元、宇树科技中标 1.24 亿人形机器人大单. [in Chinese]

[242] 艾邦机器人 (2026). 2025 年机器人年度订单 TOP10，优必选近 14 亿元登顶销冠. [in Chinese]

[243] American Psychological Association (APA Monitor) (2026). AI chatbots and digital companions are reshaping emotional connection. https://www.apa.org/monitor/2026/01-02/trends-digital-ai-relationships-emotional-connection (accessed 2026-06-25)

[244] arXiv 2511.14972 (2025). Harmful Traits of AI Companions. https://arxiv.org/html/2511.14972v1 (accessed 2026-06-25)

[245] Psychiatric Times (2025). Protecting Children From Chatbot Companions. https://www.psychiatrictimes.com/view/protecting-children-from-chatbot-companions (accessed 2026-06-25)

[246] Republic of Korea Ministry of Education / Korea Herald (2025). South Korea AI Digital Textbooks — legal status stripped. https://english.moe.go.kr/boardCnts/viewRenewal.do?boardID=265&boardSeq=102075&m=0201&s=english (accessed 2026-06-25)

[247] NVIDIA (2025). NVIDIA Cosmos: World Foundation Models for Physical AI. https://www.nvidia.com/en-us/ai/cosmos/

[248] World Economic Forum (2020). Schools of the Future: Defining New Models of Education for the Fourth Industrial Revolution. https://www.weforum.org/publications/schools-of-the-future-defining-new-models-of-education-for-the-fourth-industrial-revolution/ (accessed 2026-06-25)

[249] 国务院 (2017). 新一代人工智能发展规划（国发〔2017〕35 号）. [in Chinese] https://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm (accessed 2026-06-25)

[250] 中华人民共和国教育部 (2024). 教育强国建设规划纲要 (2024--2035); 教育信息化 2.0 行动计划 (2018). [in Chinese]

[251] 新浪财经 (海外 TMT 系列报告) (2025). 优必选 2024 年营收增长稳健，人形机器人商业化加速赋能. [in Chinese]

[252] 21 世纪经济报道 (2025). 广东机器人七剑客之一优必选交卷：2024 年营收 13 亿元，同比增长 24\%. [in Chinese]

[253] InfoQ (2026). 宇树科技靠融资还是靠卖货活着？翻完招股书，我们挖到了这 8 个关键点. [in Chinese]

[254] Robots Asia (2025). DEEP Robotics Quadrupeds: Lite3, X20, X30 Specs and Prices.

AI-SLI

June 2026