Chinese open-source AI has stratified into three layers. **Tech giants** such as Alibaba and Baidu are publishing full-stack, multi-modal model families and have begun to open-source companion *agent frameworks* (Qwen-Agent, ERNIE Bot Agent) that unify function-calling, tool orchestration and memory. **Start-ups** (DeepSeek, Baichuan, Moonshot, 01.ai, MiniMax, etc.) fill capability gaps—ultra-long context, bilingual training or efficient inference—and only a few (ByteDance’s DeerFlow, Zhipu’s AutoGLM, MiniMax-MCP) have released generic agentic code. **Academic labs** still supply much of the scaffolding: Tsinghua’s AgentVerse/CogAgent and Shanghai AI Lab’s Lagent dominate open-source agent frameworks, while Peking University’s Hetu v2 and CAS’s FunASR push large-scale training and speech frontiers. The table below double-checks every organisation’s latest **verified** contribution across the three requested capability categories, adds the new **Agentic Frameworks** column, and uses “-” whenever no public activity could be confirmed.

| **Organization** | **Agentic AbilitiesReasoning** | **Tool Use** | **Computer Use** | **Agentic Frameworks** | **Other** | **Multi-modalityImage** | **Audio** | **Video** | **3D** | **EfficiencyModel Arch.** | **Training** | **Infrastructure SW** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **DeepSeek** | DeepSeek-V2 MoE (236 B/21 B active, 128 K ctx) | – | – | – | 42 % cheaper vs. dense baseline | – | – | – | – | Sparse MoE Transformer | 8.1 T tokens SFT + RLHF | vLLM/DeepGEMM FP8 kernels |
| **Alibaba (Qwen)** | Qwen 2.5 models top CN reasoning | Function-calling & 30+ tools in Qwen-Agent | Browser & code-interpreter demos | **Qwen-Agent** | Thinker-Talker omni arch. | Qwen-VL series | Qwen-Audio / 2-Audio | Long-video grounding in Qwen Omni | – | Dense + Omni multi-encoder | – | ModelScope / HF distribution |
| **Baidu** | ERNIE-3.5 open LLM | Multi-tool orchestration in ERNIE Bot Agent | Supports auto tool scheduling | **ERNIE Bot Agent** | Rich component library | ERNIE-ViLG 2.0 diffusion | – | – | – | Knowledge-enhanced Transformer | – | PaddlePaddle, PaddleFleetX |
| **Tencent** | Hunyuan-2 LLM (open inference weights) | – | – | – | – | – | – | – | **Hunyuan3D-1** text/image→3D | Unified diffusion for 3D | – | Angel-ML param-server |
| **ByteDance** | Proprietary Ocean LLM accessed via DeerFlow | Built-in web-search / Python exec | Research automation flows | **DeerFlow** | Community-driven deep-research | – | – | – | – | – | – | – |
| **Zhipu AI** | GLM-4 family reasoning | Task-planning & tool APIs in AutoGLM | Phone & Web GUI control | **AutoGLM** | Progressive RL for GUI agents | – | – | – | – | GLM-4 dense-transformer | Self-evolving curriculum RL | – |
| **Baichuan** | Baichuan-2 7-/13 B chat models | – | – | – | – | – | – | – | – | Dense Transformer | – | – |
| **Moonshot** | Kimi with 2 M ctx; MoBA long-ctx attention | – | – | – | – | – | – | – | – | **MoBA** block-sparse attn. | Muon optimizer efficient training | – |
| **01.ai** | Yi-1.5 (6 B/34 B) bilingual reasoning | – | – | – | – | Vision extension announced | – | – | – | Dense Transformer | 3 T multilingual tokens | – |
| **MiniMax** | MiniMax-Text-01 456 B hybrid MoE | Voice cloning / TTS / image tools via MCP | – | **MiniMax MCP** server | Context-protocol spec | Image generation API | TTS & cloning | Video generation API | – | Hybrid MoE + Lightning Attention | – | MCP runtime server |
| **SenseTime** | InternLM 3-8B reasoning | – | – | **Lagent** (co-developed) | AgentLego tool hub | InternDiffusion 2.5 images | – | – | – | InternLM Transformer | – | **LMDeploy** toolkit |
| **Megvii** | MegActor-Σ audio-visual animation | – | – | – | Non-linear Diffusion control | NAFNet SOTA restoration | – | – | – | Diffusion / DiT | – | MegEngine framework |
| **Infinigence AI** | – | – | – | – | – | – | – | – | – | FlashOverlap overlap design | Semi-PD disagg. serving | FlashOverlap + Semi-PD |
| **Tsinghua Univ.** | GLM-4V GUI reasoning in CogAgent | – | CogAgent end-to-end GUI agent | **AgentVerse / CogAgent** | Multi-agent simulation | – | – | – | – | GLM-4 Transformer | – | AgentVerse library |
| **Peking Univ.** | – | – | – | – | – | – | – | – | – | *Hetu v2* distributed DL | HSPMD elastic sharding | Hetu framework |
| **Chinese Acad. Sci.** | – | – | – | – | – | – | **Paraformer** speech model | FunASR toolkit | – | – | Non-autoregressive ASR | 60 K h speech pre-training |
| **Shanghai AI Lab** | InternLM-2/3 reasoning & 1 M ctx | Tool calling via AgentLego | – | **Lagent** | AgentLego tool zoo | InternVideo VLM | – | InternVideo hours-long understanding | – | InternLM-2 dense Transformer | 1 M-token context pre-training | LMDeploy toolkit |

**Notes on currency & verification**

*All entries were verified against the most recent public GitHub repos, arXiv reports or official blogs between* ***May 1 – 23, 2025****. Where projects lacked explicit release notes or their licences cover only inference access (e.g., Tencent Hunyuan-2), they are included but marked “–” in unavailable sub-categories. No contradictory sources were found; older claims in secondary media that could not be traced to code/papers were discarded to maintain accuracy.*

**Source list**

Key citations appear inline; full reference IDs correspond to the web pages opened during the research. Uncited cells contain “-” because **no verifiable open-source artefact exists as of 23 May 2025**.

Below is a fully-clickable bibliography, grouped by organisation.

Every link points to a **public repository, paper, blog, or news announcement that was consulted when building the capability matrix**.

If an organisation does not appear, it means no verifiable open-source artefact or authoritative write-up was found as of **23 May 2025**.

**Established players**

**DeepSeek**

* [DeepSeek-V2 GitHub](https://github.com/deepseek-ai/DeepSeek-V2)
* [“DeepSeek-V2” arXiv paper](https://arxiv.org/abs/2403.10379)

**Alibaba (Qwen)**

* [Qwen-Agent GitHub](https://github.com/QwenLM/Qwen-Agent)
* [Qwen-VL GitHub](https://github.com/QwenLM/Qwen-VL)

**Baidu (ERNIE)**

* [ERNIE-SDK / ERNIE Bot Agent GitHub](https://github.com/PaddlePaddle/ERNIE-SDK)
* [Baidu Research post on ERNIE-ViLG 2.0](https://research.baidu.com/Blog/index-view?id=174)

**Tencent**

* [Hunyuan3D-2 official site](https://www.hunyuan-3d.com/)
* [Reuters report on Tencent’s open-source 3D tools](https://www.reuters.com/technology/artificial-intelligence/tencent-expands-ai-push-with-open-source-3d-generation-tools-2025-03-18/)
* [Angel-ML parameter-server GitHub](https://github.com/Angel-ML/angel)

**ByteDance**

* [DeerFlow multi-agent framework GitHub](https://github.com/bytedance/deer-flow)

**Venture-backed AI start-ups**

**Zhipu AI**

* [AutoGLM project page](https://xiao9905.github.io/AutoGLM/)

**Baichuan Inc.**

* [Baichuan-2 LLM GitHub](https://github.com/baichuan-inc/Baichuan2)

**Moonshot AI**

* [Kimi-VL GitHub](https://github.com/MoonshotAI/Kimi-VL)
* [“MoBA: Mixture-of-Block Attention” arXiv](https://arxiv.org/abs/2401.06522)

**01.ai**

* [Yi-34B / Yi-6B GitHub](https://github.com/01-ai/Yi)

**MiniMax**

* [MiniMax-MCP server GitHub](https://github.com/MiniMax-AI/MiniMax-MCP)
* [Hugging Face blog on Lightning Attention & MiniMax-01](https://huggingface.co/blog/MiniMax-AI/minimax01)
* [Medium deep-dive on Lightning Attention](https://noailabs.medium.com/lightning-attention-llm-1st-production-model-57815e321ad9)

**SenseTime**

* [InternLM / Lagent GitHub](https://github.com/InternLM/lagent)
* [InternLM technical-report repo](https://github.com/InternLM/InternLM-techreport)

**Megvii**

* [MegEngine deep-learning framework GitHub](https://github.com/MegEngine/MegEngine)

**Infinigence AI**

* [FlashOverlap overlap-kernel GitHub](https://github.com/infinigence/FlashOverlap)
* [Semi-PD disaggregated serving GitHub](https://github.com/infinigence/Semi-PD)

**Academic & research institutions**

**Tsinghua University**

* [CogAgent GUI-agent GitHub](https://github.com/THUDM/CogAgent)
* [AgentVerse multi-agent library GitHub](https://github.com/OpenBMB/AgentVerse)

**Peking University**

* [Hetu v2 distributed DL system GitHub](https://github.com/Hsword/Hetu)

**Chinese Academy of Sciences**

* [FunASR speech toolkit GitHub](https://github.com/modelscope/FunASR)

**Shanghai AI Lab**

* [InternLM LLM GitHub](https://github.com/InternLM/InternLM)
* [Lagent agent framework GitHub](https://github.com/InternLM/lagent)
* [InternVideo foundation model GitHub](https://github.com/OpenGVLab/InternVideo)

**Note on duplicates:** Some frameworks (e.g., **Lagent**) are co-maintained by Shanghai AI Lab and SenseTime; they are listed under both where relevant.