



State of AI: China

Artificial Analysis
Q2 2025

Highlights Report

Full report available to Premium Access subscribers

Artificial Analysis is a leading, and independent AI benchmarking and insights provider. We support engineers and companies to understand AI capabilities and make critical decisions about their AI strategy.

Our data, insights and publications are grounded in our comprehensive benchmarking of AI technologies and use cases. This includes everything from hourly performance testing of language model APIs to millions of votes in our crowd-sourced arenas.

Our public website, artificialanalysis.ai, is widely referenced by companies leading innovation in AI. To discuss this report, our publications, or our services, please get in touch at contact@artificialanalysis.ai.

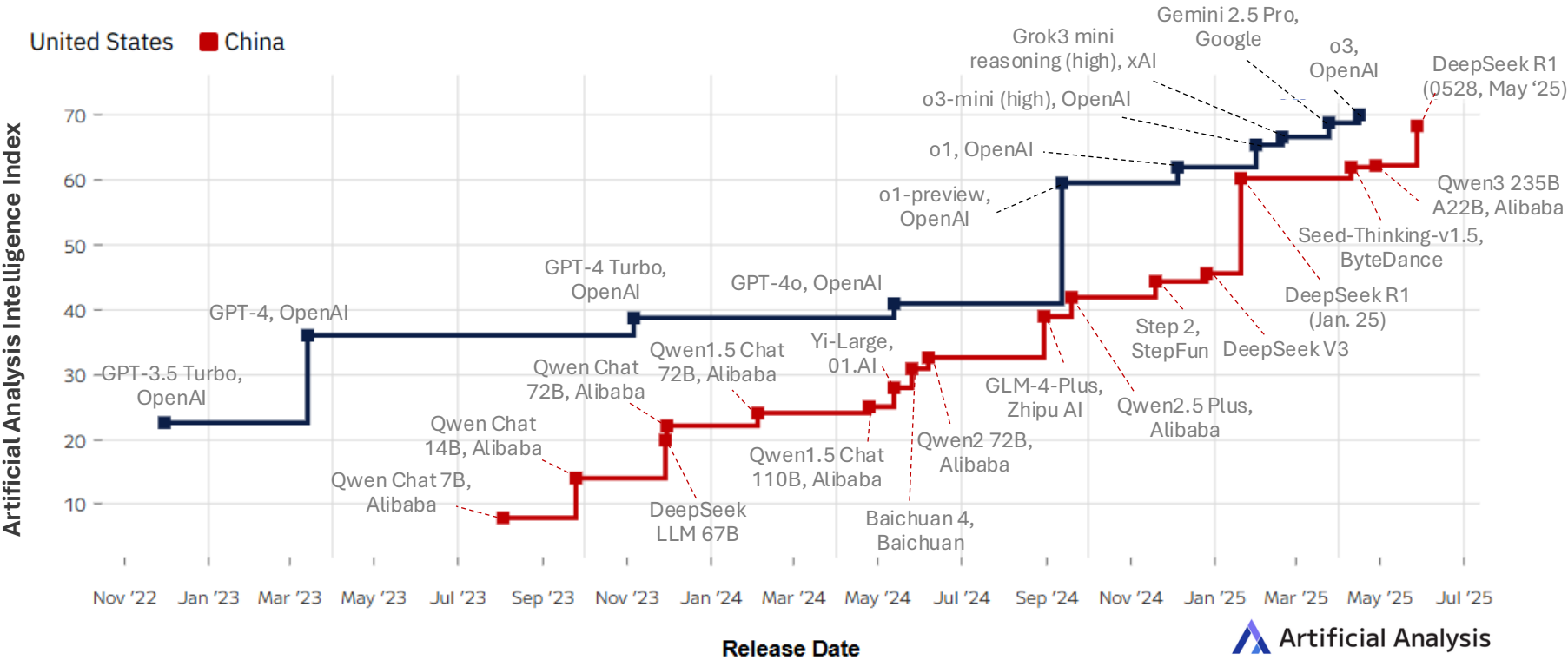




China’s leading AI labs are now closer than ever to US leaders, with the lead decreasing from more than a year to less than three months

US & China: Frontier Language Model Intelligence, Over Time

Artificial Analysis Intelligence Index incorporates 7 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME, MATH-500



Commentary

- The performance gap between US and Chinese frontier models since the release of ChatGPT in 2022 has remained persistent, but is as narrow now as it has ever been
- DeepSeek’s open weights R1 (May 2025) model leads the Chinese AI labs, while OpenAI’s o3 leads models released by US AI Labs
- DeepSeek and Alibaba have primarily driven the Chinese frontier, while advances in the US frontier have been primarily driven by OpenAI

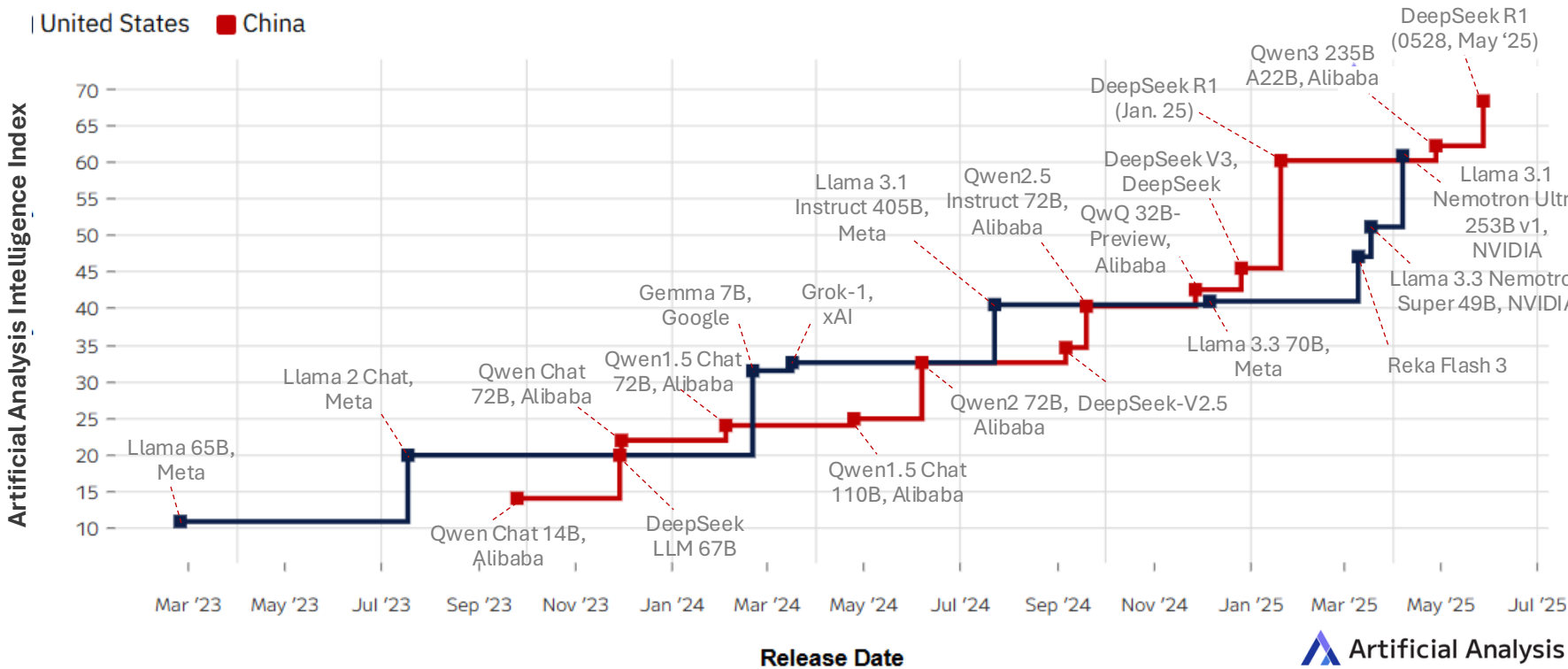
Note: Some results were estimated based on company claims and comparable results
Source: Artificial Analysis Intelligence Index



The Chinese open weights frontier surpassed the US in November 2024 with Alibaba's release of QwQ 32B Preview, R1 consolidated this lead

US & China: Open Weights Frontier Language Model Intelligence, Over Time

Artificial Analysis Intelligence Index incorporates 7 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME, MATH-500



Commentary

- **The Chinese open weights frontier surpassed the US in November 2024** with the release of QwQ 32B Preview (overtaking Meta's Llama 3.1 405B)
- **The open weights leadership of Chinese AI labs is reflective of the approach of the top Chinese AI labs to often release the weights of their flagship models.** This contrasts with the top US AI labs, which generally do not release the weights of their leading models, e.g., OpenAI, Anthropic and Google
- **China's DeepSeek R1 (January 2025) was the first open weights reasoning model to be competitive with OpenAI's o1**
- **DeepSeek's R1 0528 (May 2025) is the most intelligent open weights model currently available**

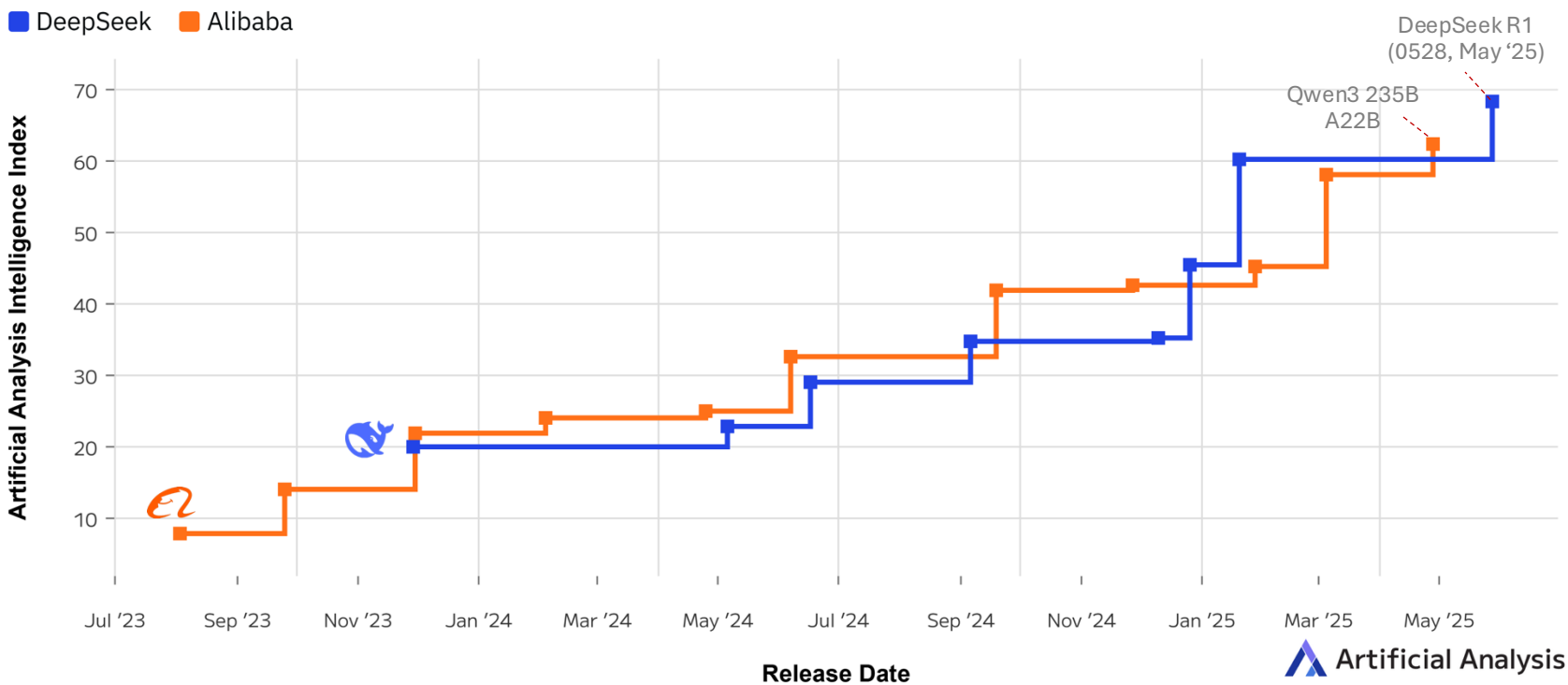
Note: Some results were estimated based on company claims and comparable results
Source: Artificial Analysis Intelligence Index



Leading Chinese AI labs DeepSeek and Alibaba have steadily released new models, with DeepSeek taking the lead in late 2024

Leading Chinese AI Labs: Language Model Intelligence, Over Time

Artificial Analysis Intelligence Index incorporates 7 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME, MATH-500



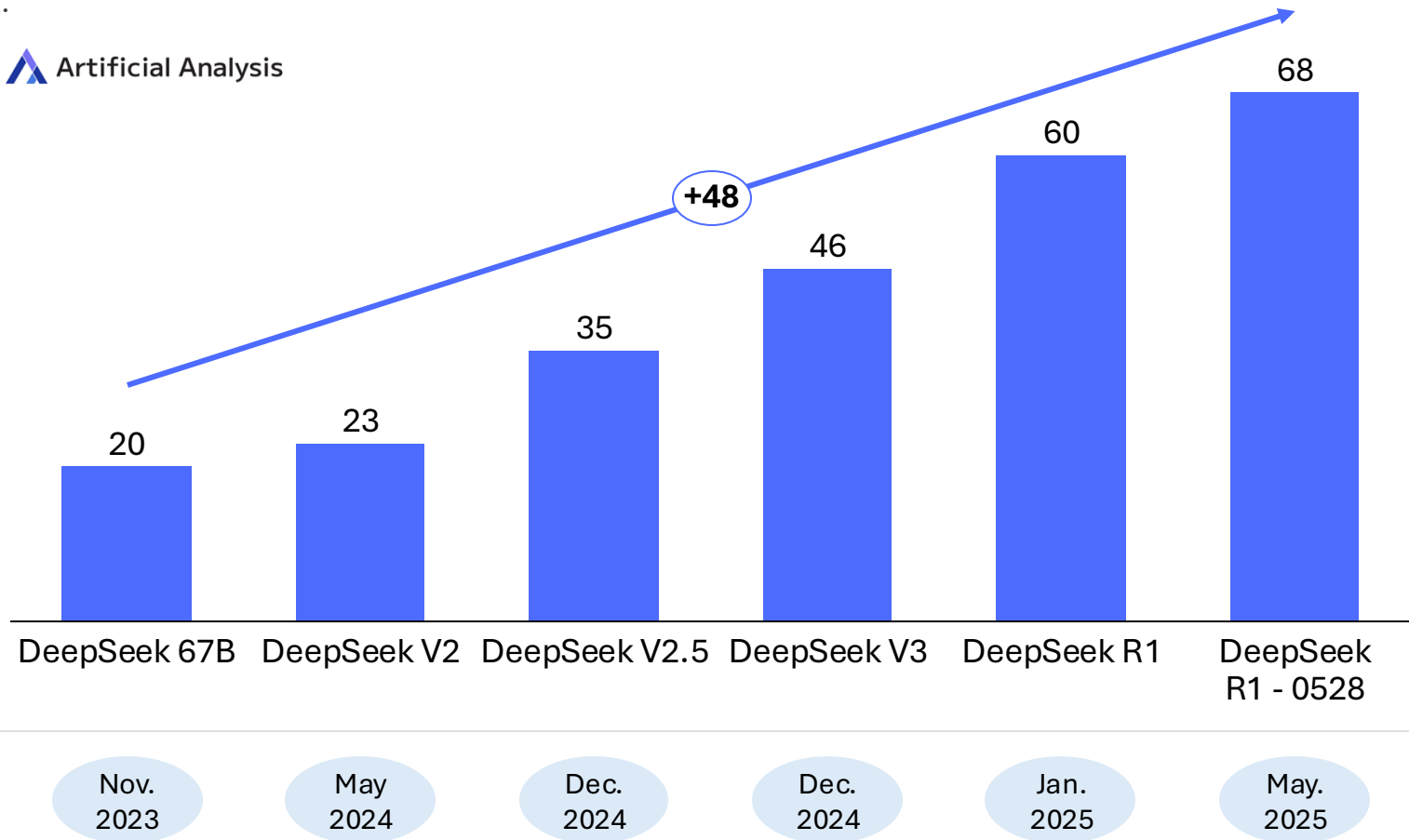


DeepSeek’s models have quickly increased in intelligence since their first public language model release in November 2023

DeepSeek’s Model Release Timeline: Artificial Analysis Intelligence Index

Only highest intelligence general-purpose DeepSeek models shown, task-specific models excluded. Artificial Analysis Intelligence Index incorporates 7 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME, MATH-500.

 Artificial Analysis



Source: Artificial Analysis Intelligence Index

Commentary

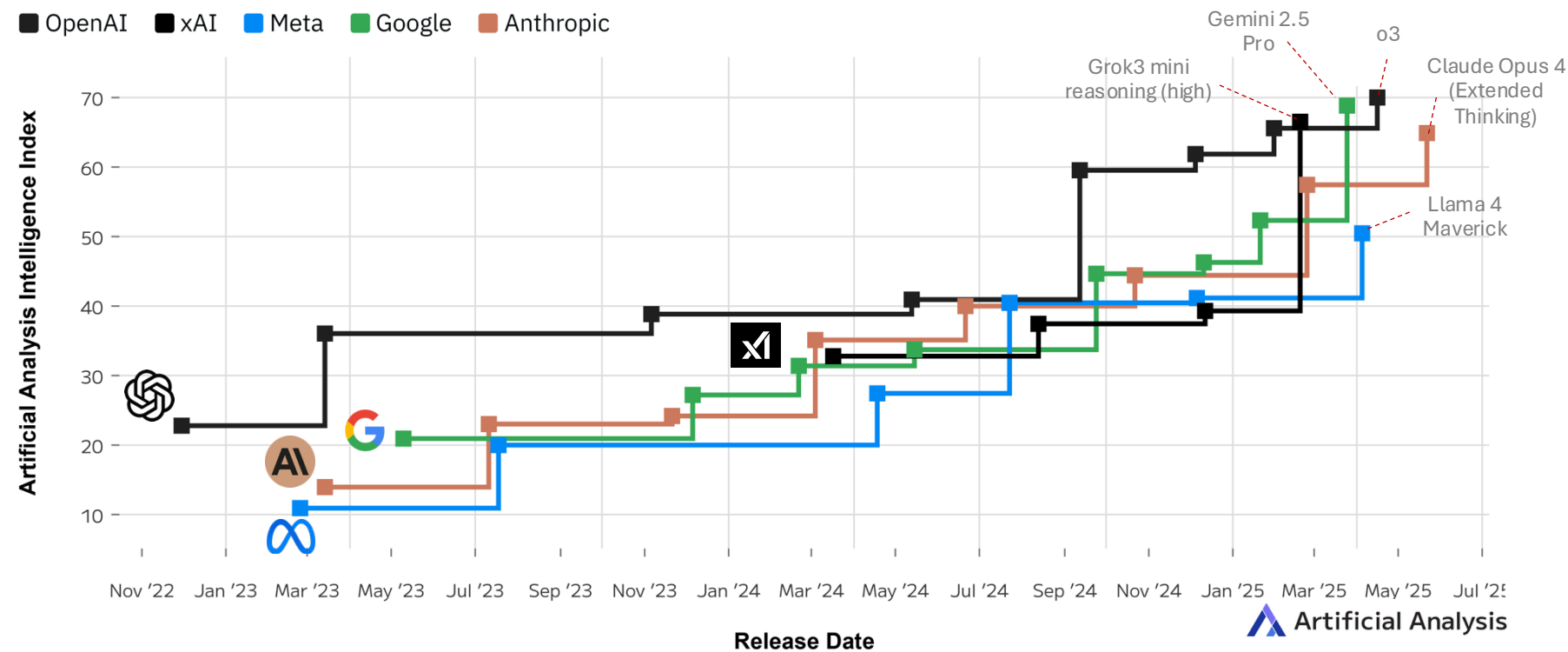
- **DeepSeek leaped over xAI, Meta and Anthropic to be tied as the world’s #2 AI lab** and the undisputed open weights leader with the release of R1-0528
- **R1-0528’s impressive intelligence uplift was a post-training update with no change to the original V3/R1 architecture** - it remains a large 671B model with 37B active parameters
- **This highlights the continually increasing importance of post-training**, particularly for reasoning models trained with reinforcement learning (RL) techniques



In the US, there are now multiple contenders for AI leadership, with OpenAI no longer in a dominant position at the frontier

Leading US AI Labs: Language Model Intelligence, Over Time

Artificial Analysis Intelligence Index incorporates 7 evaluations: MMLU-Pro, GPQA Diamond, Humanity's Last Exam, LiveCodeBench, SciCode, AIME, MATH-500



Commentary

- OpenAI has been the clear leader of the AI intelligence frontier, but its lead has diminished as other labs including Google, xAI and Anthropic have narrowed the gap
- As of May 2025, OpenAI's o3 is the most intelligent US model, and the most intelligent model overall

Note: Some results were estimated based on company claims and comparable results
Source: Artificial Analysis Intelligence Index




There are many Chinese AI players, and they categorize into 3 broad archetypes

The focus of this report, deep dive to follow.

1. “Big tech” players


Existing large technology players in China

Alibaba
NYSE: BABA (Mkt Cap: ~\$300B)




Large ecommerce player and hyperscaler (Alibaba Cloud), largest shareholder of Ant Group

ByteDance
Private




Parent company of Douyin (TikTok), Toutiao, and Doubao, the largest AI chatbot in China with ~60M monthly active users

Huawei
Private




Global telco leader, hyperscaler, and leader in domestic semiconductor manufacturing for AI chips.

Tencent
HKG: 700 (Mkt Cap: ~\$575B)



Parent company of Riot Games and WeChat, the 'all-in-one' 'super app' of China; hyperscaler with their Tencent Cloud offering

Baidu
NASDAQ: BIDU (Mkt Cap: ~\$30B)




China's largest search engine, and operator of Wenxin Yiyan, an AI chatbot with ~13M monthly active users (MAU)

2. AI startups


Smaller focussed AI companies

DeepSeek




Leading Chinese AI lab originating out of High-Flyer, an AI-focussed quantitative trading firm; best known for DeepSeek R1

Moonshot AI




China AI Tiger with 2M Chinese character context window model; owner of popular AI chatbot Kimi with ~13M MAU

Zhipu




China AI Tiger spun out from a research project at Tsinghua University in 2019; Strong commitment to open weights models

StepFun




First Chinese AI startup to develop a 1T param model; founded by Jiang Daxin (ex-Chief Scientist, Microsoft Research Asia)

MiniMax




China AI Tiger and publisher of Talkie AI app (4th most downloaded app in the US in 1H24)

01.AI



Founded by Lee Kai-Fu (former head of Google China); working closely with Alibaba to set up a “industrial large model lab”

Baichuan




China AI Tiger with a focus on medical AI models

NOT EXHAUSTIVE

3. Other companies with AI ambitions


Other technology players

Xiaomi
HKG: 1810 (Mkt Cap: ~\$165B)




China's leading consumer electronics brand. Poached Luo Fuli, DeepSeek researcher, to run AI lab and recently released MiMo-7B, a small reasoning model.

Kunlun Tech
SHE: 300418 (Mkt Cap: ~\$5.5B)




Beijing-based internet group with >300m MAUs; owner of the Opera browser; Launched SkyWork AI models and accelerators

360 Security (Qihoo 360)
SHA: 601360 (Mkt Cap: ~\$10B)




China's largest provider of Internet and mobile security. Launched the Zhiniao series of models under the 360 AI brand

iFlytek
SHE: 002230 (Mkt Cap: ~\$15B)



Leading voice AI company in China with >14,000 employees. Launched the Spark series of models

Meituan
HKG: 3690 (Mkt Cap: ~\$110B)








China's top shopping platform with >600m DAUs. Co-founder Wang Huiwen returned to lead AI efforts. Investor in frontier labs



Leading Chinese “big tech” firms are all pursuing frontier-class AI models, and several have consumer AI apps with over 100M users

In-depth Description of Chinese “big tech” Firms










































					
Description	Large ecommerce player and hyperscaler (Alibaba Cloud), largest shareholder of Ant Group	Parent company of Douyin (TikTok), Toutiao, and Doubao, the largest AI chatbot in China with ~60 monthly active users	Global telco leader, hyperscaler, and leader in domestic semiconductor manufacturing for AI chips.	Parent company of Riot Games and WeChat; hyperscaler with their Tencent Cloud offering	China’s largest search engine, and operator of Wenxin Yiyan, an AI chatbot with ~13M monthly active users (MAU)
AI Strategy (high-level)	<ul style="list-style-type: none">• Releasing leading open weights models (Qwen) and tooling (ModelScope)• Offering model inference through Alibaba Cloud to other businesses	<ul style="list-style-type: none">• Developing leading proprietary models in language, image, and video• Integrating models across modalities into consumer facing products (TikTok, Toutiao, Doubao)• Offering model inference through Volcano Engine (ByteDance’s cloud offering)	<ul style="list-style-type: none">• Building AI infrastructure from chips (Ascend 910C), systems (CloudMatrix 384)• Developing proprietary models on their own hardware (Pangu Ultra)• Offering model inference and chip lease through Huawei Cloud to other businesses	<ul style="list-style-type: none">• Focussing on developing proprietary models• Integrating models into consumer-facing chatbot (Yuanbao), and WeChat• Offering model inference through Baidu Cloud to other businesses	<ul style="list-style-type: none">• Focussing on developing proprietary models• Integrating models into consumer-facing search product, and chatbot (Wenxin Yiyan, formerly ERNIE Bot)• Offering model inference through Baidu Cloud to other businesses
Best Non-Reasoner LLM	Qwen3-235B-A22B Intelligence: 47	Doubao 1.5 Pro Intelligence: 48*	Pangu 5.0 Large ¹	Hunyuan TurboS Intelligence: 47*	ERNIE 4.5 Intelligence: 46*
Best Reasoner LLM	Qwen3-235B-A22B (Reasoning) Intelligence: 62	Seed Thinking v1.5 Intelligence: 62*	-	Hunyuan T1 Intelligence: 59*	ERNIE X1 Intelligence: 60*
Consumer Apps	Tongyi Qianwen (~150M MAU)	Doubao (~110M MAU)	Celia (~200M MAU)	Yuanbao, Yuanqi (~40M MAU)	Wenxin Yiyan, Wenxin Yige (~13M MAU)
Market Cap (US\$)	~300B	Private	Private	~\$575B	~30B

1. No publicly published benchmarks or international API available; *Estimated based on company claims and comparable results where available, not yet independently benchmarked by Artificial Analysis
Source: Artificial Analysis Intelligence Index, Crunchbase, aicpb.com, Press Search



There is a deep ecosystem of well-funded Chinese AI lab startups with competitive AI models, many have financial backing by Chinese “big tech” players

In-depth Description of Chinese Startups

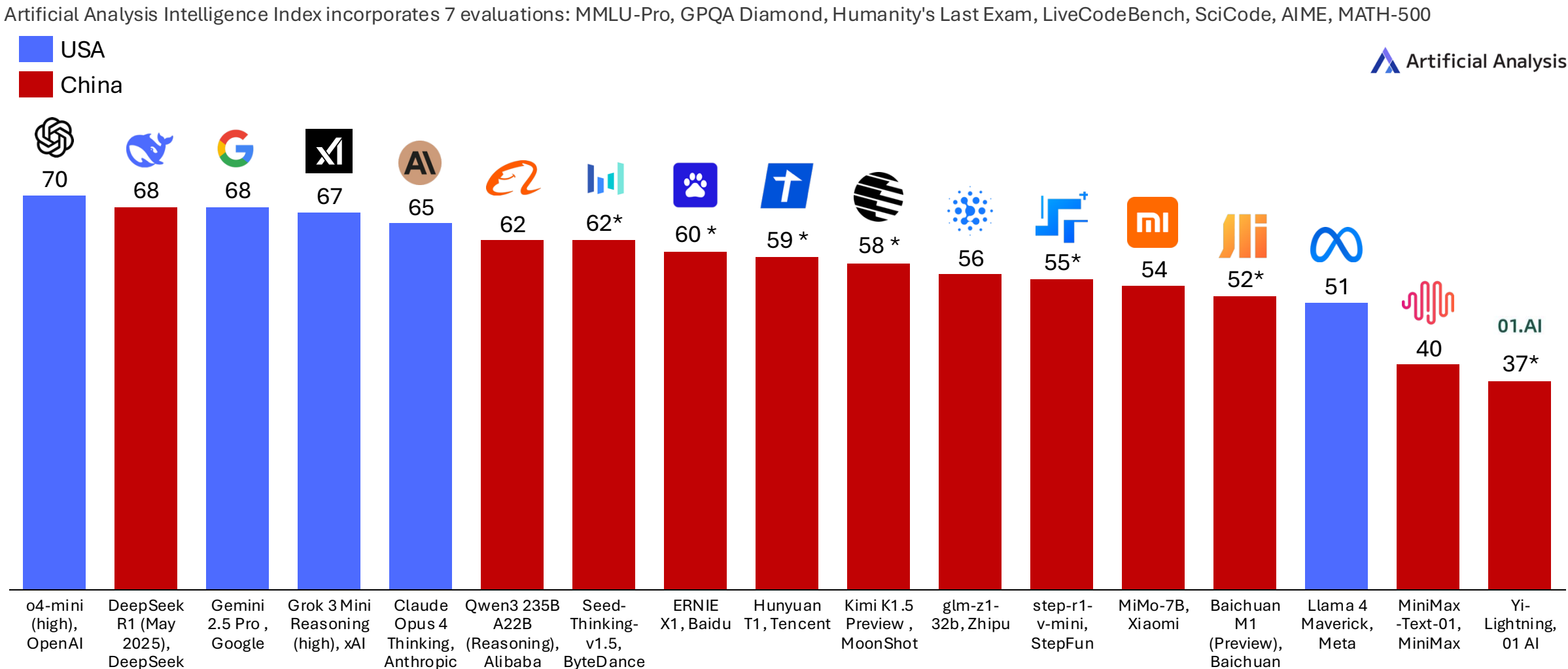
	 deepseek	 Moonshot AI	 ZHIPU-AI	 阶跃星辰 Stepfun	 MINIMAX	 零一万物 01.AI	 百川智能 BAICHUAN AI
Description	Leading Chinese AI lab originating out of High-Flyer, an AI-focussed quantitative trading firm; best known for DeepSeek R1	China AI Tiger with 2M Chinese character context window model; owner of popular AI chatbot Kimi with ~13M MAU	China AI Tiger spun out from a research project at Tsinghua University in 2019; strong commitment to open weights models	First Chinese AI startup to develop a 1T param model; founded by Jiang Daxin (ex-Chief Scientist, Microsoft Research Asia)	China AI Tiger and publisher of Talkie AI app (4 th most downloaded app in the US in 1H24)	Founded by Lee Kai-Fu (ex- Head of Google China); working closely with Alibaba on an “industrial large model lab”	China AI Tiger with a focus on medical AI models
Founded	May 2023	March 2023	June 2019	April 2023	December 2021	March 2023	April 2023
Notable investors	 幼方量化 HIGH-FLYER QUANT	 Alibaba.com  Tencent 腾讯  HONGSHAN 红杉中国  Microsoft  ZhenFund  Xiaohongshu (Rednote)  招商局集团 CHINA MERCHANTS GROUP	 Alibaba.com  Tencent 腾讯  HONGSHAN 红杉中国  QIMING VENTURE PARTNERS  Lightspeed CRINA PARTNERS  美团 Meituan  清华控股 TSINGHUA HOLDINGS 北京市国有资产经营有限责任公司  上海国有资本投资有限公司 Shanghai State-owned Capital Investment Co., Ltd.	 Tencent 腾讯  QIMING VENTURE PARTNERS  上海国有资本投资有限公司 Shanghai State-owned Capital Investment Co., Ltd.	 Alibaba.com  Tencent 腾讯  HONGSHAN 红杉中国  高榕 GAORONG VENTURES  SINO VATION VENTURES	 Alibaba.com  xiaomi  顺为 SHUNWEI  SUSV  清华控股 TSINGHUA HOLDINGS 北京市国有资产经营有限责任公司	 Alibaba.com  Tencent 腾讯  xiaomi  顺为 SHUNWEI  清华控股 TSINGHUA HOLDINGS 北京市国有资产经营有限责任公司
Best Non-Reasoner LLM	DeepSeek V3 Intelligence: 52	Moonshot v1 Intelligence: 30*	GLM-4-32B Intelligence: 46	Step-2 Intelligence: 44	MiniMax-Text-01 Intelligence: 40	Yi-Lightning Intelligence: 37*	Baichuan 4 Intelligence: 31*
Best Reasoner LLM	DeepSeek R1 Intelligence: 68	Kimi K1.5 Intelligence: 58*	GLM-Z1-32B Intelligence: 56	Step-R1-V-Mini Intelligence: 55*	-	-	Baichuan M1 Intelligence: 52*
Consumer Apps	DeepSeek Chat (~100M MAU)	Kimi (~25M MAU)	ChatGLM (~9M MAU)	Yuewen, PopDuck (<1M MAU)	Talkie AI (~35M MAU)	YiChat (<1M MAU)	Bai Xiaoying (<1M MAU)
Funding raised (US\$)	Undisclosed	~1.6B	~1.1B	Undisclosed	~0.85B	~0.2B	~1.0B

*Estimated based on company claims and comparable results where available, not yet independently benchmarked by Artificial Analysis
Source: Artificial Analysis Intelligence Index, Crunchbase, aicpb.com, Press Search



While US AI labs maintain an overall lead in the intelligence frontier, Chinese AI labs are no longer far behind

Chinese and US Labs' Leading Language Models (non-exhaustive)



*Estimated based on company claims and comparable results where available, not yet independently benchmarked by Artificial Analysis
Source: Artificial Analysis Intelligence Index

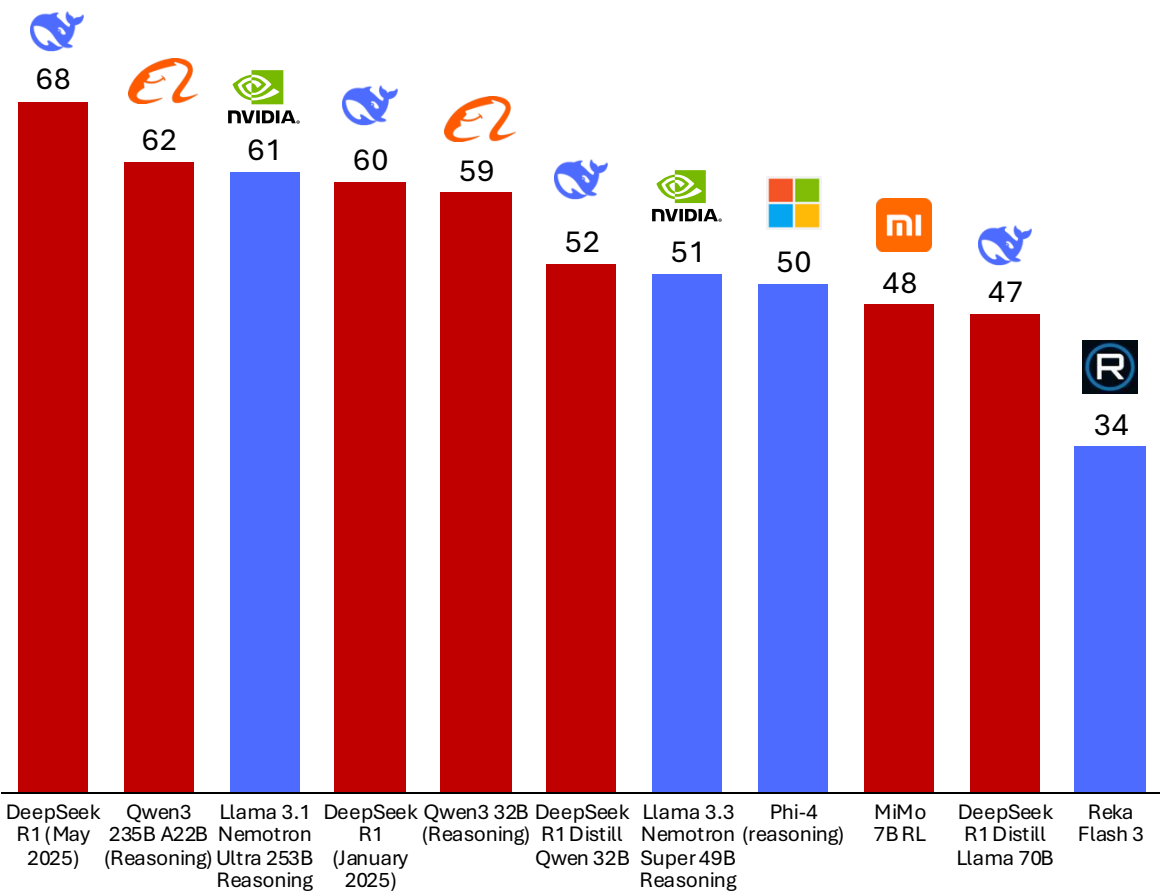


Today's open weights frontier is led by DeepSeek (both reasoning and non-reasoning models)

Reasoning: Open Weights Language Models, by Country

Artificial Analysis Intelligence Index, leading open weights reasoning models

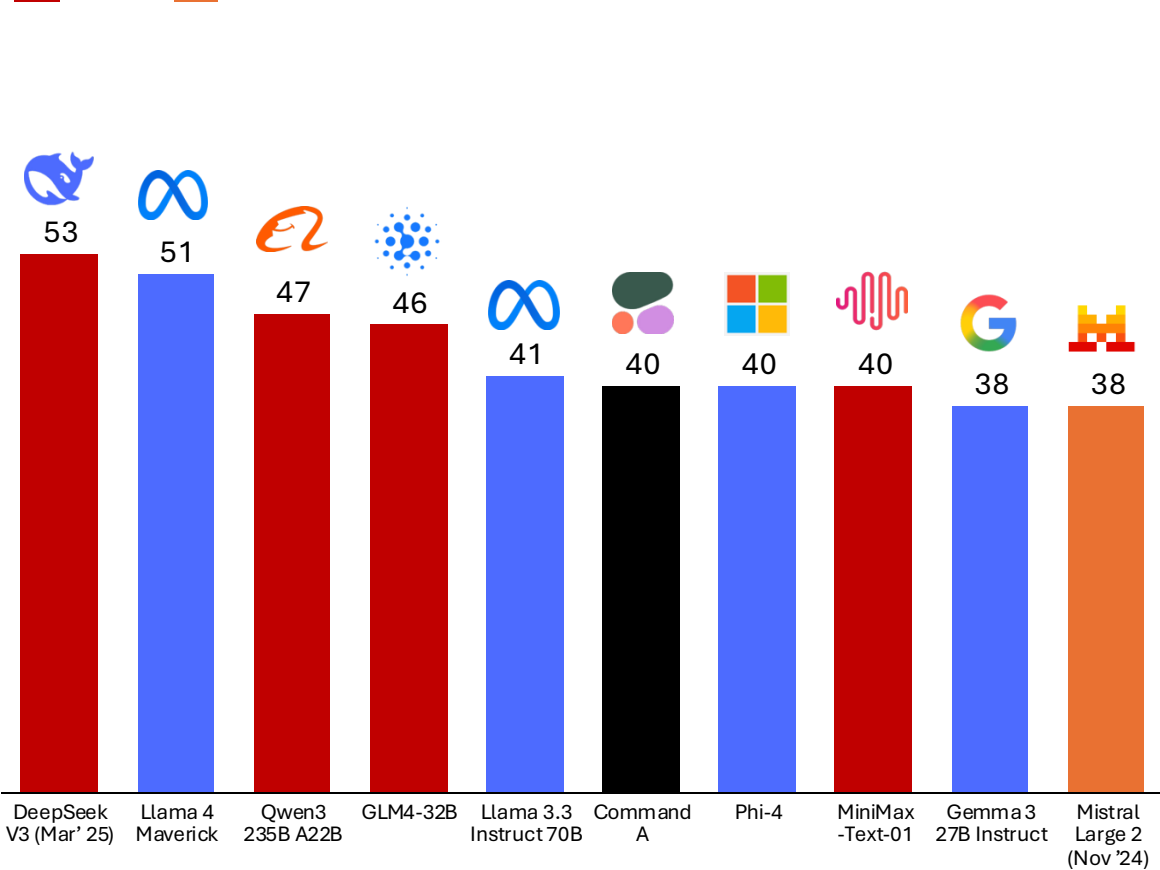
USA China Artificial Analysis



Non-reasoning: Open Weights Language Models, by Country

Artificial Analysis Intelligence Index, leading open weights non-reasoning models

USA Canada China France Artificial Analysis



Source: Artificial Analysis Intelligence Index



Chinese companies play across every major modern AI modality – including language models and all types of media generation models

Leading model for each modality, by company

No Model Research Model Only Publicly Released Model

Includes models released since January 2024, Not exhaustive

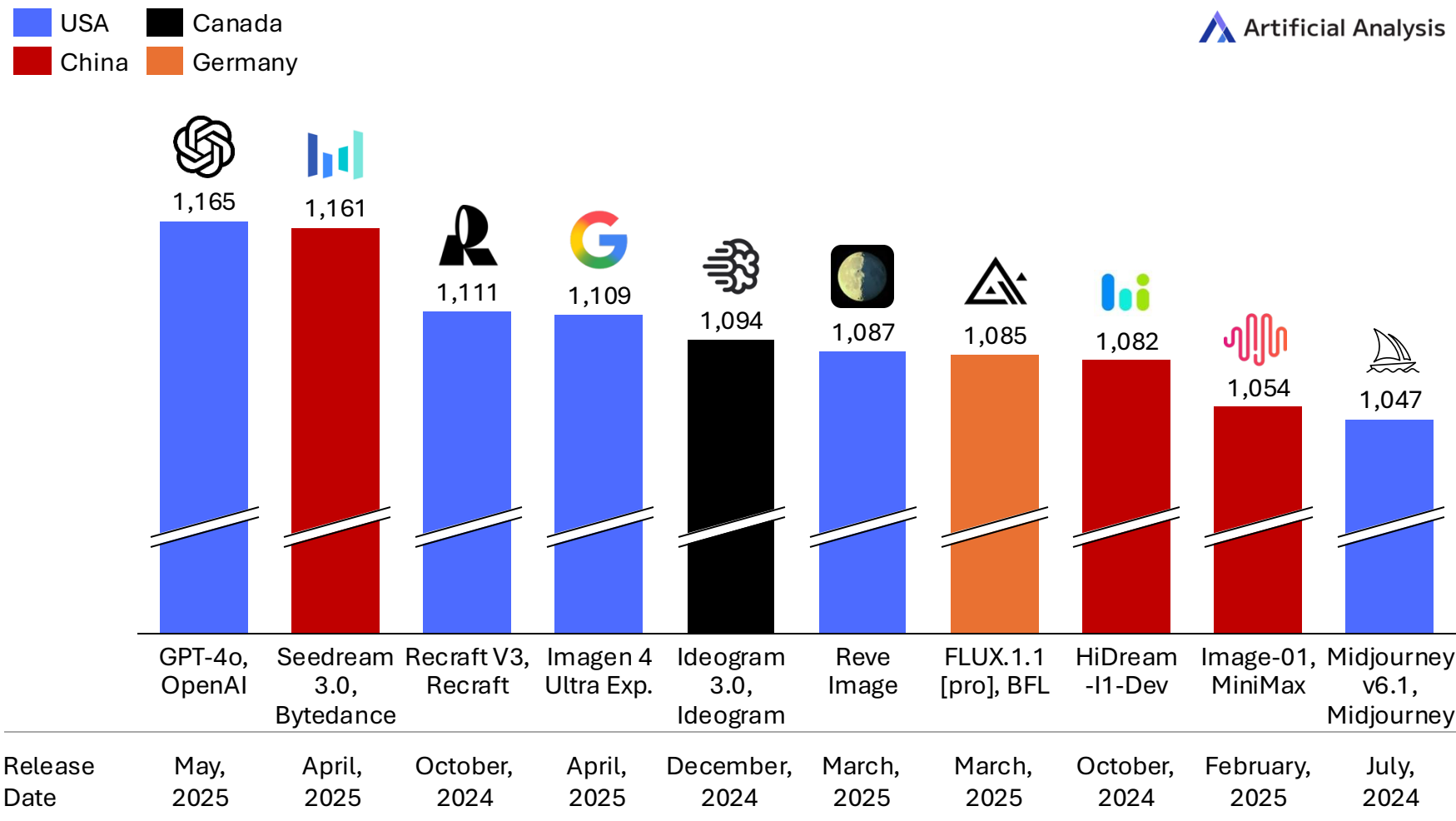
	Company	Language, Reasoning	Language, Non-Reasoning	Text to Speech	Speech to Speech	Image Gen.	Video Gen.	3D Gen.
“Big tech” Players	Alibaba	Qwen3 235B A23B		Qwen2.5 Omni		Wan 2.1		LHM
	Baidu	ERNIE X1	ERNIE 4.5	-	-	I-RAG	UniVG	Xiling
	ByteDance	Seed-Thinking-v1.5	Doubao-1.5-pro	Seed-TTS	Doubao-1.5-pro	Seedream 3.0	Seaweed-7B	ImageDream
	Huawei	PanGu Ultra MoE	Pangu 5.0	-	-	PixArt	-	-
	Tencent	Hunyuan-T1	Hunyuan-TurboS	-	-	HunyuanDiT	HunyuanVideo	Hunyuan3D
AI Startups	DeepSeek	DeepSeek R1	DeepSeek V3	-	-	Janus Pro 7B	-	-
	MoonShot AI	Kimi k1.5	Moonshot-v1	-	Kimi-Audio	-	-	-
	Zhipu	Glm-z1-32b	Glm-4-32b	-	-	CogView4	CogVideo	-
	Stepfun	Step-R1-V-Mini	Step-2	Step-audio		Step-1x	StepVideo T2V	Step1X 3D
	MiniMax	-	MiniMax-Text-01	Speech-02 HD	-	Image-01	Video-01	-
	01.AI	-	Yi-Lightning	-	-	-	-	-
	Baichuan	Baichuan-M1	Baichuan 4	Baichuan-Audio		-	-	-



Text to Image model quality between US and China reached parity in 2025

Leading Text to Image Models, Artificial Analysis Arena ELOs

Text to Image model ELOs from the Artificial Analysis Image Arena



Commentary

- China and the US are at effective parity in text to image generation, with OpenAI’s GPT-4o maintaining a 4 point ELO lead over ByteDance’s Seedream 3.0
- Text to Image leaders hail from a wide diversity of geographies, unlike language models where the US and China dominate, driven by smaller startups with a specific focus on image models
- Some older models continue to perform well against recent releases, indicating a slower innovation rate than in language models, with Recraft V3 (October 2024) remaining notably competitive despite newer entrants

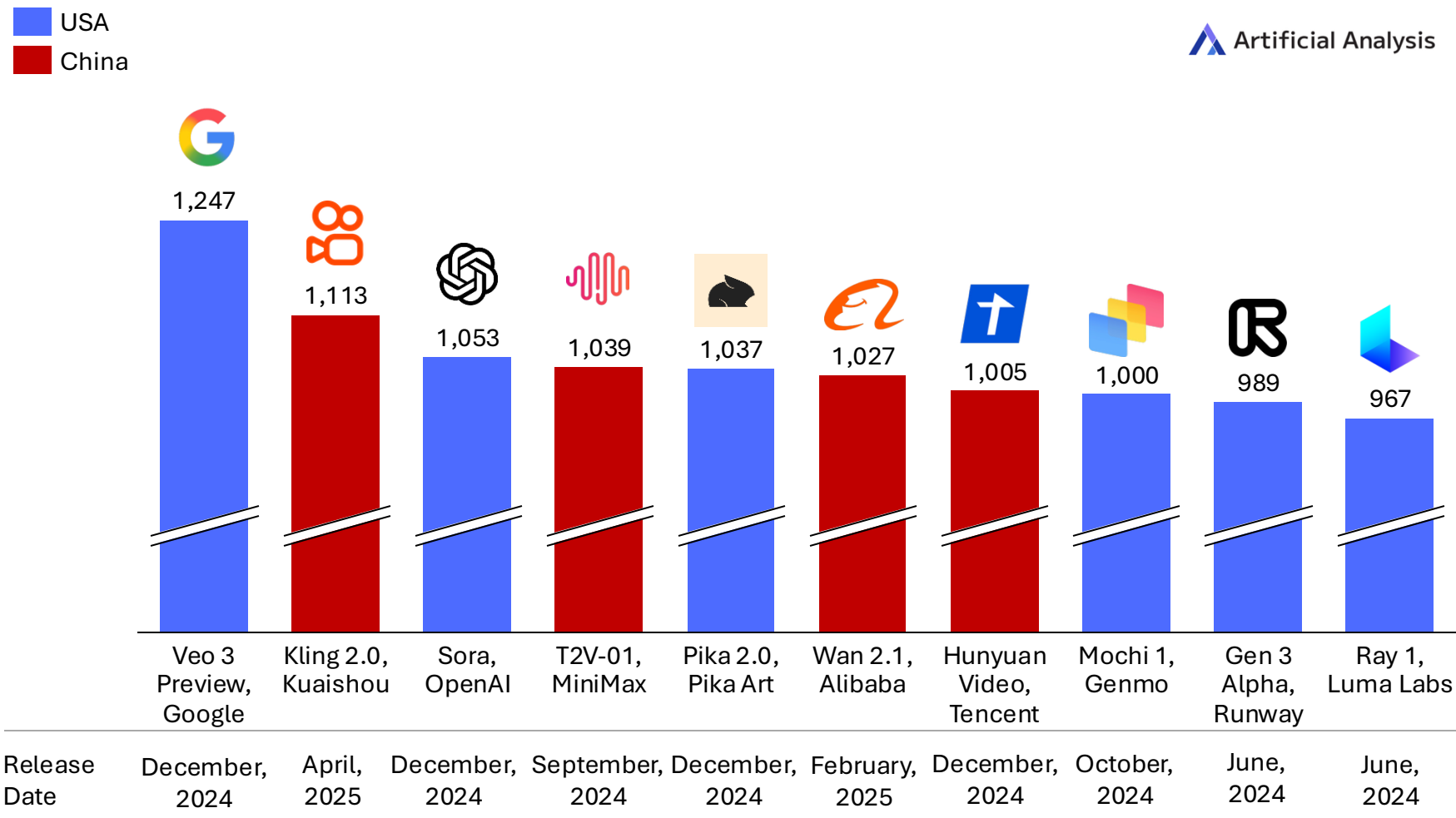
Source: Artificial Analysis Media Generation Arena Leaderboards



Text to Video is led by US and Chinese labs, with Google’s Veo 3 leading the category by a wide margin

Leading Text to Video Models, Artificial Analysis Video Arena ELOs

Text to Video model ELOs from the Artificial Analysis Video Arena



Source: Artificial Analysis Media Generation Arena Leaderboards

Commentary

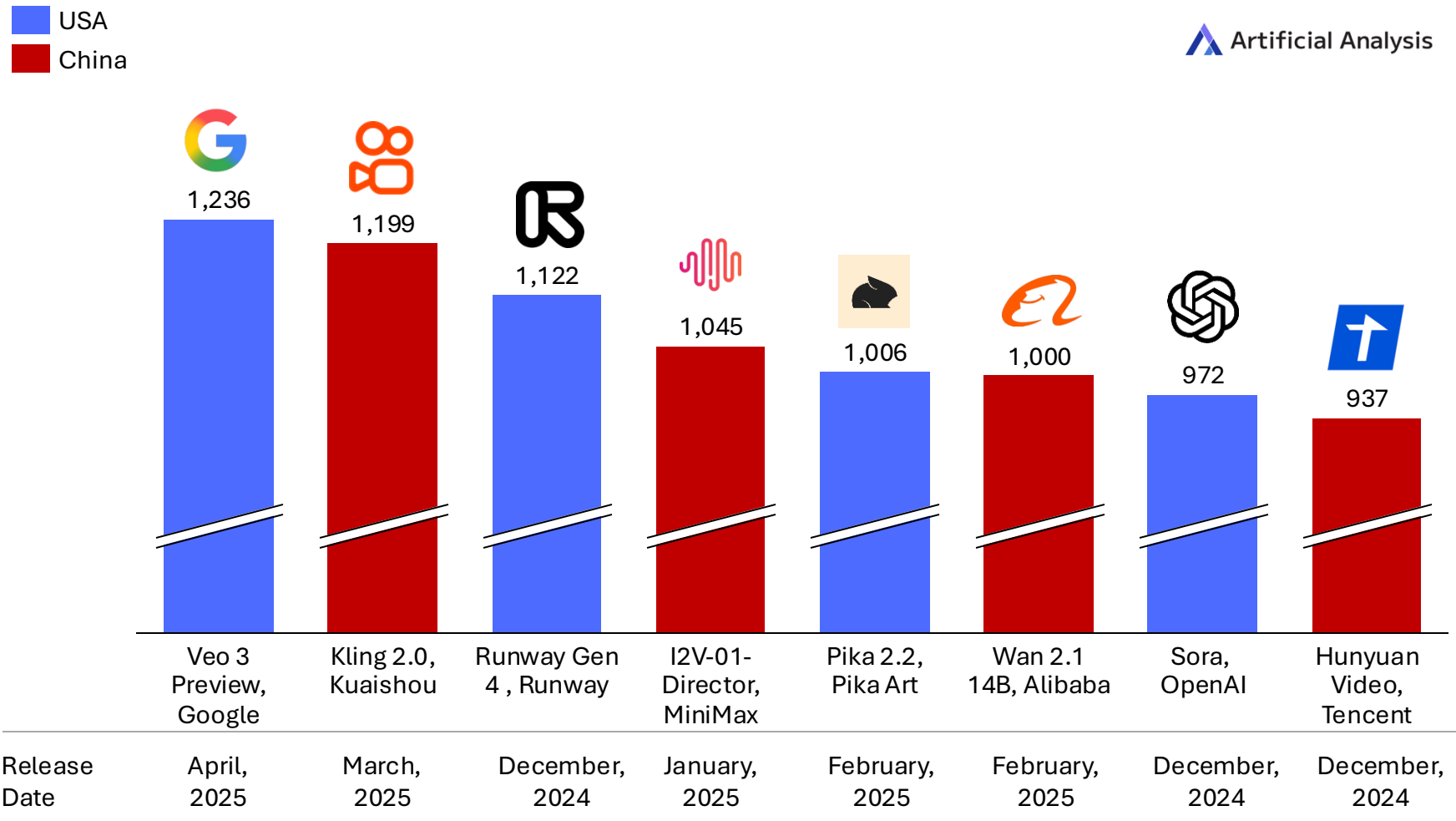
- **The US maintains a lead in text to video models**, with Google's Veo 3 Preview achieving a significantly higher ELO in the Artificial Analysis Video Arena than Kuaishou's Kling 2.0
- **Chinese companies are widely competitive in the text to video space**, with labs such as MiniMax and Alibaba performing at comparable levels to OpenAI and Pika



Image to Video is led by the US, with Google’s Veo 3 recently surpassing both Kuaishou’s Kling 2.0 and Runway Gen 4

Leading Image to Video Model ELOs, Artificial Analysis Video Arena

Text to Video model ELOs from the Artificial Analysis Video Arena



Source: Artificial Analysis Media Generation Arena Leaderboards

Commentary

- **Google’s Veo 3 Preview maintains a small lead in Image to Video generation quality**, surpassing Kuaishou’s Kling 2.0 and Runway’s Gen 4 in the Artificial Analysis Video Arena
- **The quality of video generation models can vary dramatically between input modalities**, with some models like Sora excelling at text to video generation while underperforming in image to video tasks
- **A recent trend has emerged of video generation models being released with Image to Video capabilities only**, as seen with Runway Gen 4 and Pika 2.2



Artificial Analysis

hello@artificialanalysis.ai

<https://artificialanalysis.ai/>

Legal notice:

Copyright © 2025 Artificial Analysis, Inc. All rights reserved.

This document, including any data, analysis, and insights contained herein, is provided by Artificial Analysis for informational purposes only. The information is based on data collected through various sources, including but not limited to first party benchmarking and surveys conducted on our website. While Artificial Analysis strives to ensure the accuracy and reliability of the information, it is provided “as is” and may not be complete or up to date. The content should not be construed as professional advice, and recipients are encouraged to conduct their own research and analysis before making any decisions based on this information. By accessing or using this document, you agree to be bound by Artificial Analysis’s Terms of Service, available on our website.