

[Open in app ↗](#)**Medium**

Search



Write

**Member-only story**

DeepSeek Is Chinese But Its AI Models Are From Another Planet

OpenAI and the US are in deep trouble

Alberto Romero · [Follow](#)

13 min read · Jan 22, 2025

5.3K

156



...



Yutong Liu / Better Images of AI / Joining the Table / CC-BY 4.0

As my #1 prediction for AI in 2025 I wrote this: “The geopolitical risk discourse (democracy vs authoritarianism) will overshadow the existential risk discourse (humans vs AI).” DeepSeek is the reason why.

I’ve covered news about DeepSeek ten times since December 4, 2023, in this newsletter. But I’d bet you a free yearly subscription that you didn’t notice the name as something worth watching. It was just another unknown AI startup. Making more mediocre models. Surely not “at the level of OpenAI or Google” as I wrote a month ago. “At least not yet,” I added, confident. Turns out I was delusional. Or bad at assessing the determination of these people. Or perhaps I was right back then and they’re damn *quick*.

Whatever the case, DeepSeek, the silent startup, will now be known. Not because it’s Chinese — that too — but because the models they’re building are outstanding. And because they’re open source. And because they’re cheap. *Damn cheap.*

Yesterday, January 20, 2025, they announced and released DeepSeek-R1, their first reasoning model (from now on R1; try it here, use the “deepthink” option). R1 is akin to OpenAI o1, which was released on December 5, 2024. We’re talking about a one-month delay — a brief window, intriguingly, between leading closed labs and the open-source community. A brief window, *critically*, between the United States and China.

From my prediction, you may think I saw this coming. Well, I didn’t see it coming *this soon*. Wasn’t OpenAI half a year ahead of the rest of the US AI labs? And more than one year ahead of Chinese companies like Alibaba or Tencent?

Others saw it coming better. In a Washington Post opinion piece published in July 2024, OpenAI CEO, Sam Altman argued that a “democratic vision for AI must prevail over an authoritarian one.” And warned, “The United States currently has a lead in AI development, but continued leadership is far from guaranteed.” And reminded us that “the People’s Republic of China has said that it aims to become the global leader in AI by 2030.” Yet I bet even he’s surprised by DeepSeek.

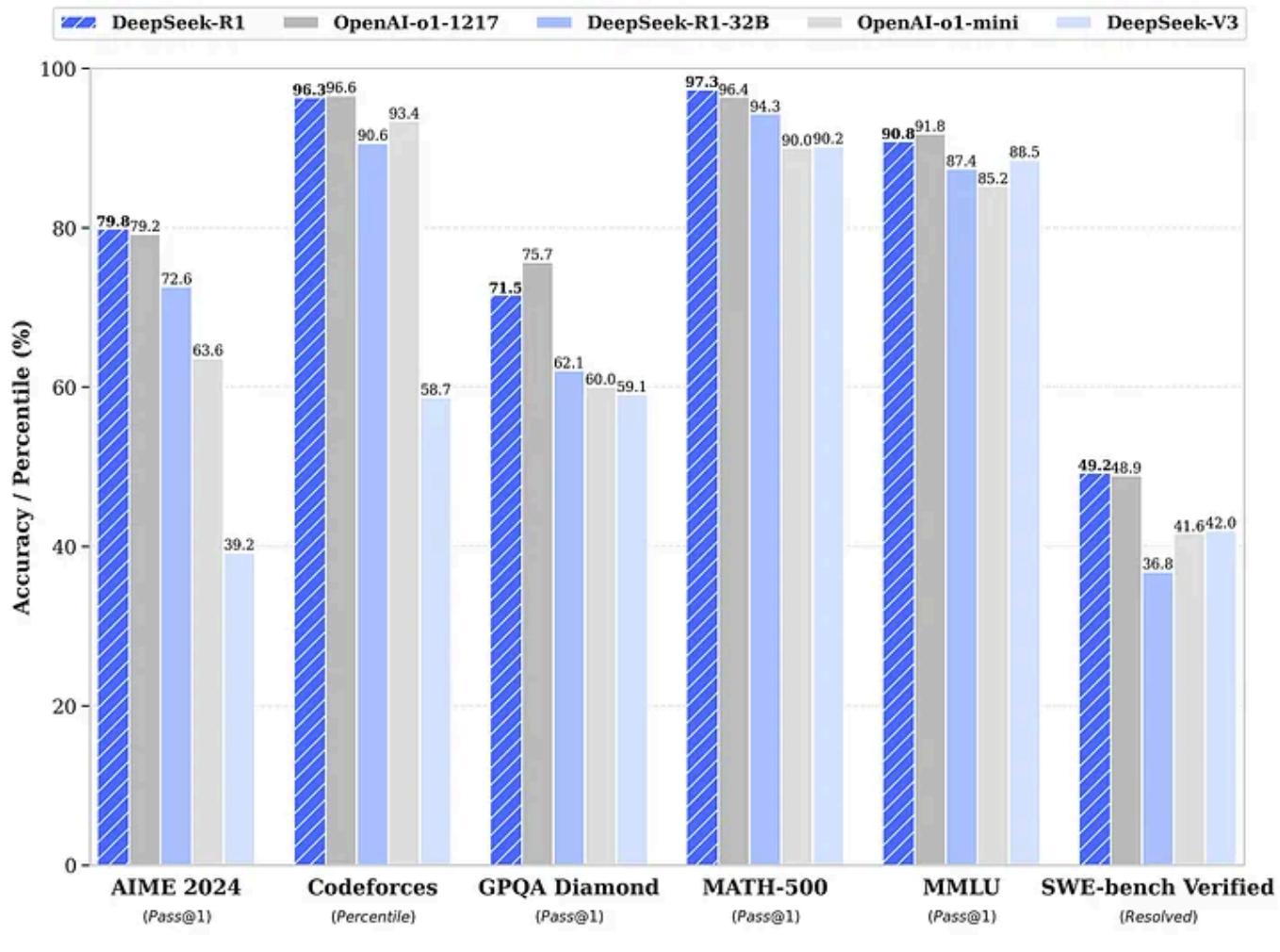
It is rather ironic that OpenAI still keeps its frontier research behind closed doors — even from US peers so the authoritarian excuse no longer works — whereas DeepSeek has given the entire world access to R1. There are too many readings here to untangle this apparent contradiction and I know too little about Chinese foreign policy to comment on them. So I will just highlight three questions I consider relevant before getting into R1, the AI model, which is the part I can safely write about:

- Does China aim to overtake the United States in the race toward AGI, or are they moving at the necessary pace to capitalize on American companies’ slipstream?
- Is DeepSeek open-sourcing its models to collaborate with the international AI ecosystem or is it a means to attract attention to their prowess before closing down (either for business or geopolitical reasons)?
- How did they build a model so good, so quickly and so cheaply; do they know something American AI labs are missing?

Now that we’ve got the geopolitical aspect of the whole thing out of the way we can concentrate on what really matters: bar charts.

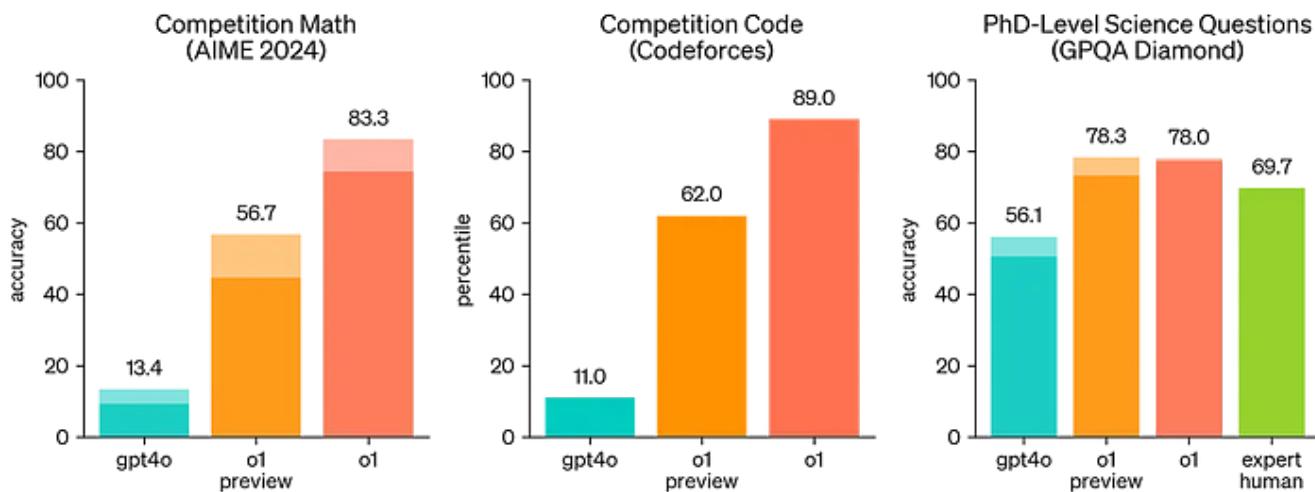
II. How good is R1 compared to o1?

DeepSeek shared a one-on-one comparison between R1 and o1 on six relevant benchmarks (e.g. GPQA Diamond and SWE-bench Verified) and other alternative tests (e.g. Codeforces and AIME). (The list is missing ARC-AGI and FrontierMath but OpenAI had an “unfair” advantage so I understand if other labs choose to ignore those.):



[Source](#)

R1 and o1 are basically the same model as measured by these evaluations, with the largest difference being a small 4-percentage point gap on GPQA Diamond (71.5% vs 75.7%). For comparison, here's o1 vs GPT-4o (OpenAI's best non-reasoning model) on three of those benchmarks (Anthropic's Claude 3.5 and Google DeepMind's Gemini 2.0 are in the ballpark of GPT-4o):



[Source](#)

That's an 85% gap between R1 and GPT-4o on Codeforces (coding). 75% gap on AIME (math). And a 15% gap on GPQA Diamond (science). In short: DeepSeek has created a model comparable to OpenAI's best model, which is also the best among US labs (it's worth noting that OpenAI has already announced o3, much better than o1, but it's yet to start rolling it out over the next months). DeepSeek is on the podium and by open-sourcing R1 it's giving away the prize money.

So let's talk about what else they're giving us because R1 is just one out of *eight* different models that DeepSeek has released and open-sourced. There's R1-Zero which will give us plenty to talk about. What separates R1 and R1-Zero is that the latter wasn't guided by human-labeled data in its post-training phase. In other words, DeepSeek let it figure out by itself how to do reasoning. More on that soon.

Then there are six other models created by training weaker base models (Qwen and Llama) on R1-distilled data. For those of you who don't know, distillation is the process by which a large powerful model "teaches" a smaller less powerful model with synthetic data. Distillation was a centerpiece in my speculative article on GPT-5. The fact that the R1-distilled

models are much better than the original ones is further evidence in favor of my hypothesis: GPT-5 exists and is being used internally for distillation. Here's a table with the results for the six distilled models:

	AIME 2024 pass@1	AIME 2024 cons@64	MATH- 500 pass@1	GPQA Diamond pass@1	LiveCodeBench pass@1	CodeForces rating
GPT-4o-0513	9.3	13.4	74.6	49.9	32.9	759.0
Claude-3.5-Sonnet-1022	16.0	26.7	78.3	65.0	38.9	717.0
o1-mini	63.6	80.0	90.0	60.0	53.8	1820.0
QwQ-32B	44.0	60.0	90.6	54.5	41.9	1316.0
DeepSeek-R1-Distill-Qwen-1.5B	28.9	52.7	83.9	33.8	16.9	954.0
DeepSeek-R1-Distill-Qwen-7B	55.5	83.3	92.8	49.1	37.6	1189.0
DeepSeek-R1-Distill-Qwen-14B	69.7	80.0	93.9	59.1	53.1	1481.0
DeepSeek-R1-Distill-Qwen-32B	72.6	83.3	94.3	62.1	57.2	1691.0
DeepSeek-R1-Distill-Llama-8B	50.4	80.0	89.1	49.0	39.6	1205.0
DeepSeek-R1-Distill-Llama-70B	70.0	86.7	94.5	65.2	57.5	1633.0

[Source](#)

When an AI company releases multiple models, the most powerful one often steals the spotlight so let me tell you what this means: A R1-distilled Qwen-14B — which is a 14 billion parameter model, *12x smaller* than GPT-3 from 2020 — is as good as OpenAI o1-mini and much better than GPT-4o or Claude Sonnet 3.5, the best non-reasoning models. That's incredible. Distillation improves weak models *so much* that it makes no sense to post-train them ever again. Just go mine your large model. And without increasing inference costs.

Talking about costs, somehow DeepSeek has managed to build R1 at 5-10% of the cost of o1 (and that's being charitable with OpenAI's input-output pricing). That's like getting a smartphone similar to a Google Pixel or Apple iPhone that costs \$1000+ for \$50 from Xiaomi or Huawei. With the same

features and quality. How? Are they running at a loss? Did they find a way to make these models incredibly cheap that OpenAI and Google ignore? Are they copying Meta's approach to make the models a commodity? Too many open questions. I guess OpenAI would prefer closed ones.

So to sum up: R1 is a top reasoning model, open source, and can distill weak models into powerful ones. All of that at a fraction of the cost of comparable models. And it is Chinese in origin. Go figure.

If I were writing about an OpenAI model I'd have to end the post here because they only give us demos and benchmarks. DeepSeek, however, also published a detailed technical report. Neither OpenAI, Google, nor Anthropic has given us something like this. So who are our friends again? It's unambiguously hilarious that it's a Chinese company doing the work OpenAI was named to do. But enough with the jokes. It's time to open the paper. It is a precious document. Let's review the parts I find more interesting.

III. What if AI didn't need us humans?

We already saw how good is R1. What's the deal with R1-Zero? DeepSeek's approach to R1 and R1-Zero is reminiscent of DeepMind's approach to AlphaGo and AlphaGo Zero (quite a few parallelisms there, perhaps OpenAI was never DeepSeek's inspiration after all).

Let me get a bit technical here (not much) to explain the difference between R1 and R1-Zero. Both are comprised of a pre-training stage (tons of data from the web) and a post-training stage. The former is shared (both R1 and R1-Zero are based on DeepSeek-V3). The latter is what changes.

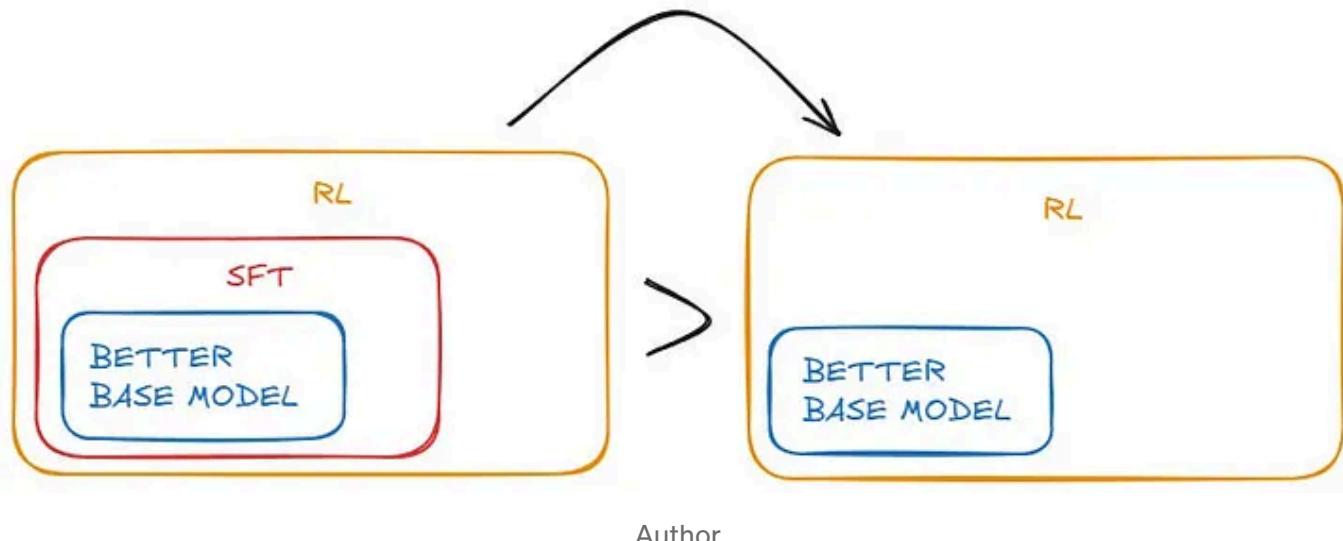
After pre-training, R1 was given a small quantity of high-quality human examples (supervised fine-tuning, SFT). That's what you normally do to get a chat model (ChatGPT) from a base model (out-of-the-box GPT-4) but in a much larger quantity. DeepSeek wanted to keep SFT at a minimum. Then, to make R1 better at reasoning, they added a layer of reinforcement learning (RL). Simple RL, nothing fancy like MCTS or PRM (don't look up those acronyms). They also allowed it to think at inference time (that's the now famous test-time compute, TTC, scaling laws that OpenAI inaugurated with o1-preview). Once you add the RL and TTC then you have something similar to o1. That's R1. R1-Zero is the same thing but without SFT.

Now without jargon: R1 surfed the web endlessly (pre-training) then read a reasoning manual made by humans (SFT), and finally did some self-experimentation (RL + TTC). R1-Zero, in contrast, didn't read any manuals. They pre-trained R1-Zero on tons of web data and immediately after they sent it to the RL phase: "Now go figure out how to reason yourself." That's it.

DeepMind did something similar to go from AlphaGo to AlphaGo Zero in 2016–2017. AlphaGo learned to play Go by knowing the rules and learning from millions of human matches but then, a year later, decided to teach AlphaGo Zero without any human data, just the rules. And it *destroyed* AlphaGo. Unfortunately, open-ended reasoning has proven harder than Go; R1-Zero is slightly worse than R1 and has some issues like poor readability (besides, both still rely heavily on vast amounts of human-created data in their base model — a far cry from an AI capable of rebuilding human civilization using nothing more than the laws of physics).

Here's what R1 vs R1-Zero looks like:

Doing SFT on the base model and then RL remains a better approach than just RL ($R1 > R1\text{-Zero}$)



But still, the relative success of R1-Zero is impressive. DeepSeek says this about the model:

The findings reveal that RL empowers DeepSeek-R1-Zero to attain robust reasoning capabilities without the need for any supervised fine-tuning data. This is a noteworthy achievement, as it underscores the model's ability to learn and generalize effectively through RL alone.

As far as we know, OpenAI has not tried this approach (they use a more complicated RL algorithm). Here are the benchmark results compared to o1-preview:

Model	AIME 2024		MATH-500	GPQA Diamond	LiveCode Bench	CodeForces
	pass@1	cons@64	pass@1	pass@1	pass@1	rating
OpenAI-o1-mini	63.6	80.0	90.0	60.0	53.8	1820
OpenAI-o1-0912	74.4	83.3	94.8	77.3	63.4	1843
DeepSeek-R1-Zero	71.0	86.7	95.9	73.3	50.0	1444

Table 2 | Comparison of DeepSeek-R1-Zero and OpenAI o1 models on reasoning-related benchmarks.

[Source](#)

But, what if it worked better? What if you could get much better results on reasoning models by showing them the entire internet and then telling them to figure out how to think with simple RL, without using SFT human data? What if — bear with me here — you didn't even need the *pre-training phase* at all? I imagine this is possible in principle (in principle it could be possible to recreate the entirety of human civilization from the laws of physics but we're not here to write an Asimov novel).

Instead of showing Zero-type models millions of examples of human language and human reasoning, why not teach them the basic rules of logic, deduction, induction, fallacies, cognitive biases, the scientific method, and general philosophical inquiry and let them discover better ways of thinking than humans could never come up with?

No human can play chess like AlphaZero. When DeepMind showed it off, human chess grandmasters' first reaction was to compare it with other AI engines like Stockfish. Soon, they acknowledged it played more like a human; beautifully, with an idiosyncratic style. I heard someone say that AlphaZero was like the silicon reincarnation of former World Chess Champion, Mikhail Tal: bold, imaginative, and full of surprising sacrifices that somehow won him so many games.

I imagine it would be harder to build such an AI program for math, science, and reasoning than chess or Go, but it shouldn't be impossible: An inhumanly smart yet uncannily humane reasoning machine. That's what DeepSeek attempted with R1-Zero and almost achieved. I'm ready for the next chapter of the story.

But let's speculate a bit more here, you know I like to do that. What if instead of becoming more human, Zero-type models get weirder as they get better?

When DeepSeek trained R1-Zero they found it *hard to read* the responses of the model. It started to mix languages. This reminds me of DeepMind again. AlphaGo Zero learned to play Go better than AlphaGo but also *weirder to human eyes*. In the end, AlphaGo had learned *from us* but AlphaGo Zero had to discover its own ways through self-play. It didn't have our data so it didn't have our flaws. More importantly, it didn't have our manners either.

Questions emerge from this: are there inhuman ways to reason about the world that are more efficient than ours? Will more intelligent AIs get not only more intelligent but *increasingly indecipherable* to us?

I believe the answer is yes: As AI gets smarter it goes through two differentiated phases. First, it gets uncannily close to human idiosyncrasy and displays emergent behaviors that resemble human “reflection” and “the exploration of alternative approaches to problem-solving,” as DeepSeek researchers say about R1-Zero. Here’s an example of “aha moment” from the paper:

Question: If $a > 1$, then the sum of the real solutions of $\sqrt{a - \sqrt{a+x}} = x$ is equal to

Response: <think>

To solve the equation $\sqrt{a - \sqrt{a+x}} = x$, let's start by squaring both ...

$$(\sqrt{a - \sqrt{a+x}})^2 = x^2 \implies a - \sqrt{a+x} = x^2.$$

Rearrange to isolate the inner square root term:

$$(a - x^2)^2 = a + x \implies a^2 - 2ax^2 + (x^2)^2 = a + x \implies x^4 - 2ax^2 - x + (a^2 - a) = 0$$

...

Wait, wait. Wait. That's an aha moment I can flag here.

Let's reevaluate this step-by-step to identify if the correct sum can be ...

We started with the equation:

$$\sqrt{a - \sqrt{a+x}} = x$$

First, let's square both sides:

$$a - \sqrt{a+x} = x^2 \implies \sqrt{a+x} = a - x^2$$

Next, I could square both sides again, treating the equation: ...

...

Table 3 | An interesting “aha moment” of an intermediate version of DeepSeek-R1-Zero. The model learns to rethink using an anthropomorphic tone. This is also an aha moment for us, allowing us to witness the power and beauty of reinforcement learning.

[Source](#)

But eventually, as AI's intelligence goes beyond what we can fathom, it gets weird; further from what makes sense to us, much like AlphaGo Zero did. It's like a comet on a long elliptical orbit, briefly meeting us in the Solar System before vanishing forever into the infinite depths of the cosmos.

I find the idea that the human way is the best way of thinking hard to defend. We're simply navigating our own flaws (the need to survive), limitations (the sequential nature of language), and cognitive blindspots (am I really smarter than everyone else, or am I just fooling myself?) There could be better ways. Unintelligible alien ways.

Perhaps OpenAI concealed o1's chain of thought not just for competitive reasons but because they arrived at a dark realization: it would be unsettling for us to witness an AI leap from English to other languages mid-sentence, then to symbols, and finally to what seems like gibberish, only to land on the

correct answer; “What the hell happened? How did you find that answer? I didn’t understand anything!!”

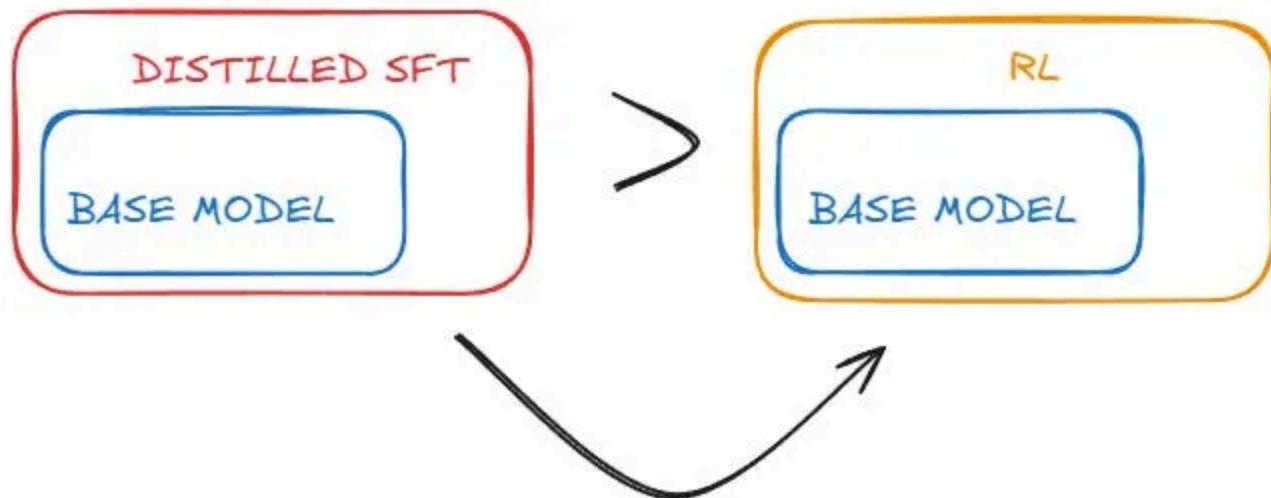
Believe me, you don’t want to look directly into the mind of an entity beyond yourself. You don’t want to shock yourself to death. I’m feeling shivers down my spine.

IV. Better base models + distillation + RL wins

Anyway, science fiction aside, there’s something else that caught my attention. DeepSeek explains in straightforward terms *what worked and what didn’t work* to create R1, R1-Zero, and the distilled models. It’s everything in there. All the secrets. Three other conclusions stand out besides what I already explained.

First, doing distilled SFT from a strong model to improve a weaker model is more fruitful than doing just RL on the weaker model. In other words, to improve a smaller, weaker model you shouldn’t apply the same approach you use to build the larger model but instead use the larger one as a teacher:

Using Qwen2.5-32B (Qwen, 2024b) as the base model, direct distillation from DeepSeek-R1 outperforms applying RL on it. This demonstrates that the reasoning patterns discovered by larger base models are crucial for improving reasoning capabilities. . . . distilling more powerful models into smaller ones yields excellent results, whereas smaller models relying on the large-scale RL mentioned in this paper require enormous computational power and may not even achieve the performance of distillation.

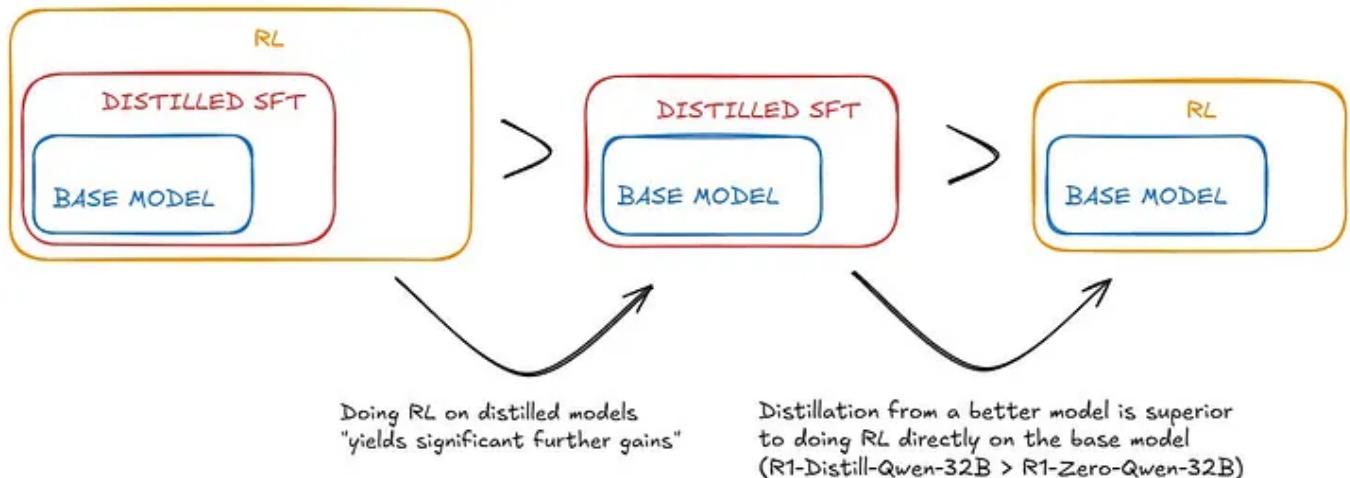


Distillation from a better model is superior to doing RL directly on the base model
 $(R1\text{-Distill-Qwen-32B} > R1\text{-Zero-Qwen-32B})$

Author

The second conclusion is the natural continuation: doing RL on smaller models is still useful.

... we found that applying RL to these distilled models yields significant further gains. We believe this warrants further exploration and therefore present only the results of the simple SFT-distilled models here.



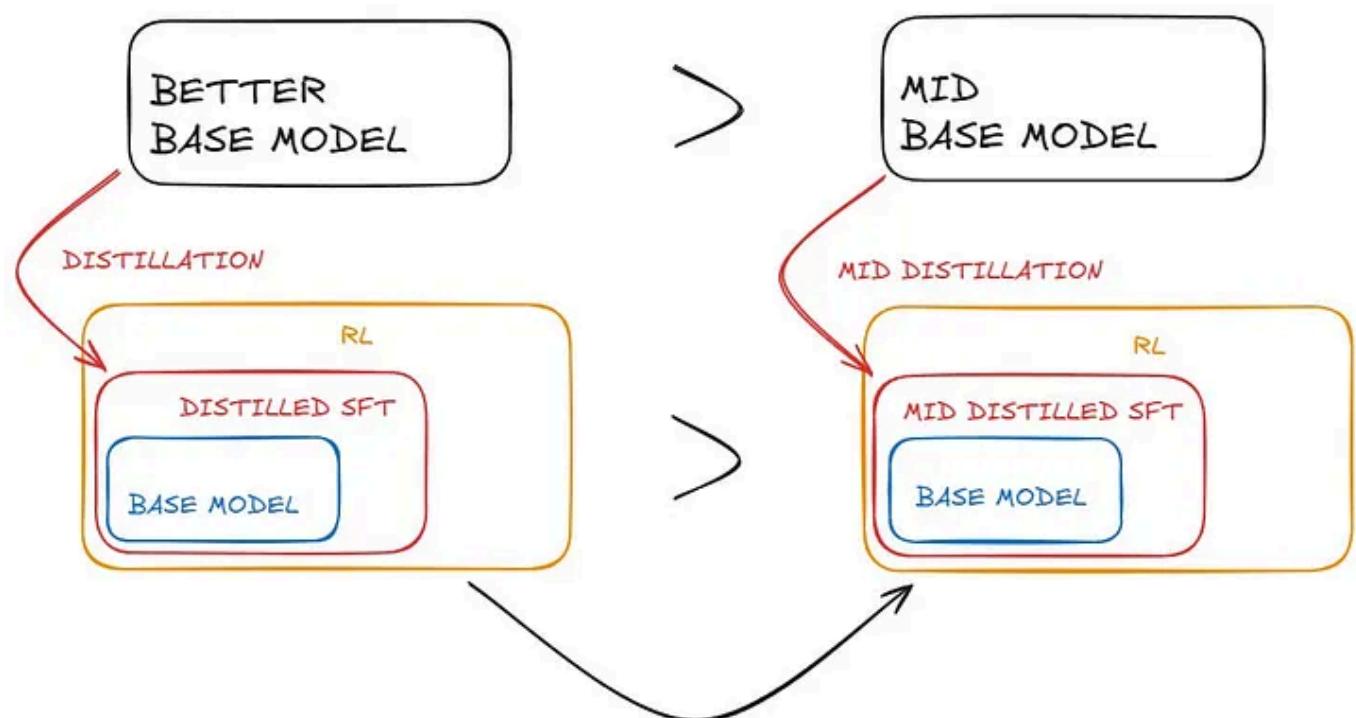
Doing RL on distilled models
 "yields significant further gains"

Distillation from a better model is superior
 to doing RL directly on the base model
 $(R1\text{-Distill-Qwen-32B} > R1\text{-Zero-Qwen-32B})$

Author

They finally conclude that to raise the floor of capability you still need to keep making the base models better. Both R1 and R1-Zero are based on DeepSeek-V3 but eventually, DeepSeek will have to train V4, V5, and so on (that's what costs tons of money). Likewise, it won't be enough for OpenAI to use GPT-5 to keep improving the o-series. At some point they will have to train a GPT-6:

... while distillation strategies are both economical and effective, advancing beyond the boundaries of intelligence may still require more powerful base models and larger-scale reinforcement learning.

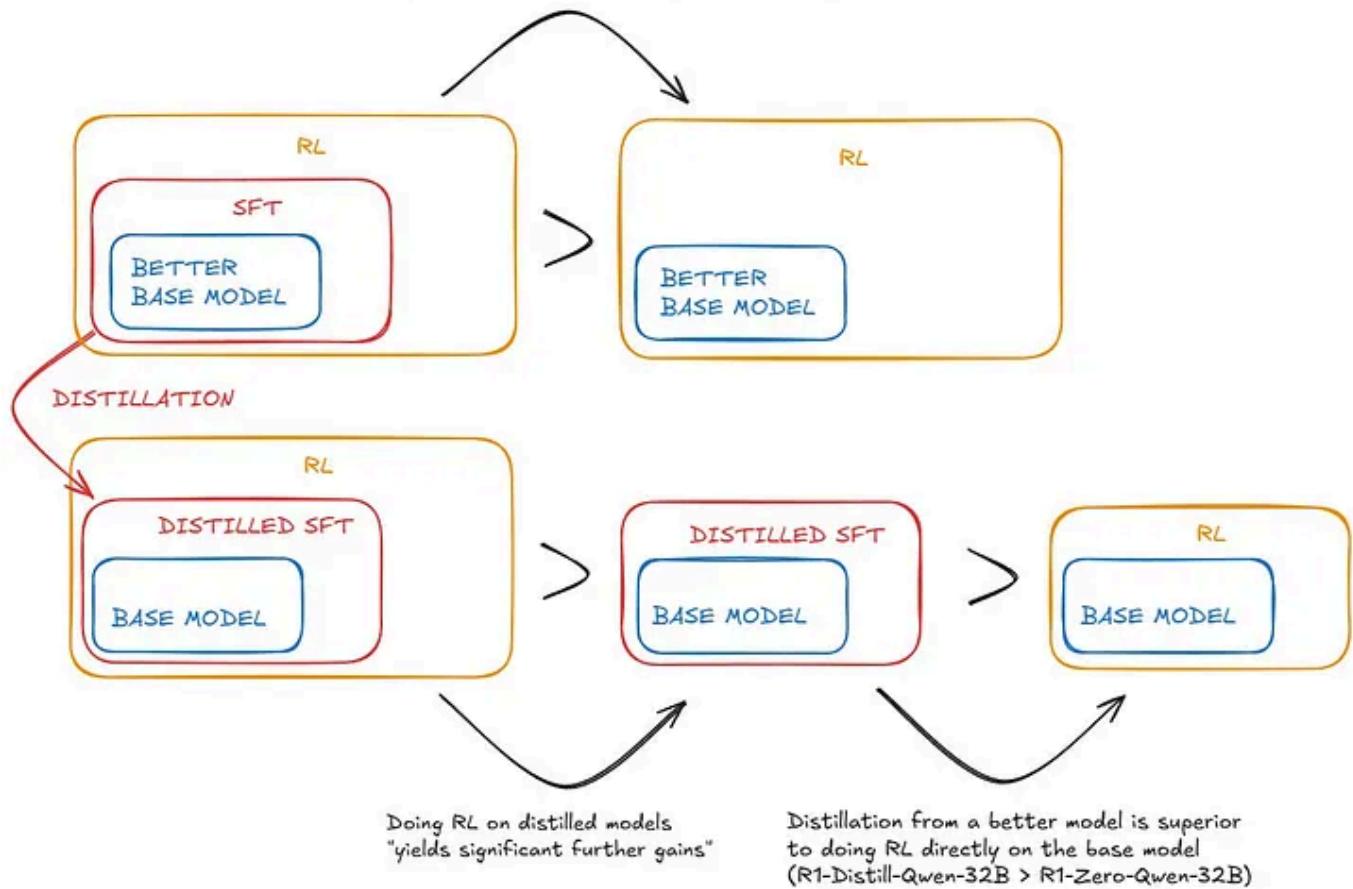


Training better base models is always useful whether to have better models or to do better distillation

Author

The whole story, in diagram form, can be summarized by this:

Doing SFT on the base model and then RL remains a better approach than just RL (R1 > R1-Zero)



Author

To DeepSeek I can only say: Congrats.

To OpenAI and the US, I have to say: Wake up.

Join The Algorithmic Bridge, a blog about AI that's actually about people.

(Join now at 20% off forever!)



Written by Alberto Romero

43K Followers · 138 Following

Follow



AI & Tech | Weekly AI Newsletter: <https://thealgorithmicbridge.substack.com/> |
Contact: alber.romgar at gmail dot com

Responses (156)



What are your thoughts?

Respond



David Anderson he/him

Jan 23

...

DeepSeek's commitment to open-source development, as evidenced by releasing R1 under the MIT license, challenges the commercial strategies of U.S. AI firms by offering similar capabilities for free. This move not only accelerates AI innovation but also fosters a more collaborative global AI community.

310

6 replies

[Reply](#)



Techyana Exploration

Jan 24

...

This article really piqued my interest, especially the description of DeepSeek's AI models as almost "alien" in their approach. It's fascinating to think about how a different cultural lens, paired with innovative technological paradigms, can lead... [more](#)

311

1 reply

[Reply](#)



David C Hager MD, MS

Jan 23

...

I'm hoping I'm wrong/off-base, but I worry about highly attractive honeypots that get leveraged to achieve other outcomes, e.g. via backdoors, Trojans, model inversion attacks, sensitive information inference, automated social engineering, etc.

Here's... [more](#)



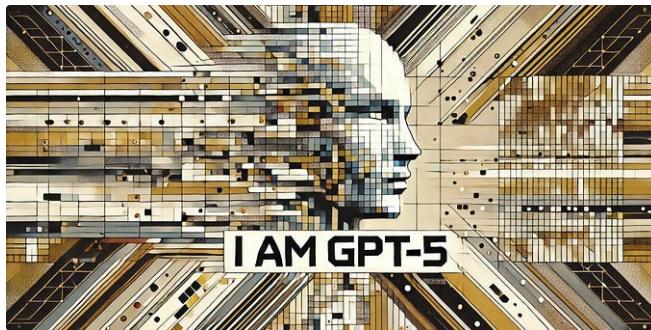
278



12 replies

[See all responses](#)

More from Alberto Romero

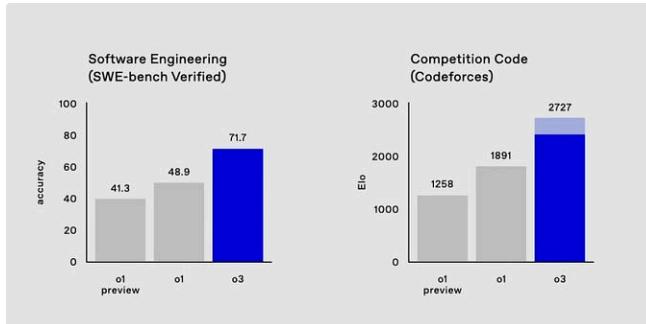


 Alberto Romero

This Rumor About GPT-5 Changes Everything

Let's start the year on an exciting note

★ Jan 16 ⚡ 995 🗣 30



 Alberto Romero

OpenAI o3 Model Is a Message From the Future: Update All You...

Incredible, a miracle, more than just a better state-of-the-art AI model

★ Dec 22, 2024 ⚡ 1K 🗣 24



Rank	Article Title	Action
1	DeepSeek - AI Assistant	Open
2	ChatGPT	Open

 Alberto Romero

7 Implications of DeepSeek's Victory Over American AI...

The Chinese Whale has caused a tsunami in the US

★ 5d ago ⚡ 1.1K 🗣 38



 Alberto Romero

How Meta Plans To Crank the Dopamine Machine With Infinite ...

Breaking: tech companies have figured out a new way to upend humanity

★ Jan 3 ⚡ 896 🗣 21



[See all from Alberto Romero](#)

Recommended from Medium



 In Age of Awareness by Cliff Berg

They Know a Collapse Is Coming

The CIO of Goldman Sachs has said that in the next year, companies at the forefront will...

Jan 21  11.3K  565



 **ReplyGuy**

The AI That finds the best places to mention your product online

ReplyGuy finds the **perfect conversations** to mention your product and drafts suggested replies.

[Try it free](#)

How It Works FAQ Pricing Blog Login

I'm looking for a way to transcribe text to speech in Chrome so I can listen to news articles, any suggestions? Replied

I also like listening to my news, specifically, book alike. I've tried several tools but my favorite is AnySpeech - fast and has great voices Replied

I finished my first novel and would like to release an audiobook version. Where did you find a narrator? Replied

Normal AI generates terrible audiobooks. I considered having AI do it? I used AnySpeech for my last audiobook and it came out great Replied

 Mohit Vaswani

6 AI Agents That Are So Good, They Feel Illegal

AI agents are the future because they can replace all the manual work with automation...

Jan 11  2.1K  81



Lists



AI Regulation

6 stories · 679 saves



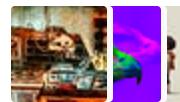
Generative AI Recommended Reading

52 stories · 1630 saves



ChatGPT prompts

51 stories · 2527 saves



What is ChatGPT?

9 stories · 499 saves

Free Apps

Paid Apps



1 DeepSeek - AI Assistant
Intelligent AI Assistant

Open

2 ChatGPT
The official app by OpenAI

Open
 Alberto Romero

7 Implications of DeepSeek's Victory Over American AI...

The Chinese Whale has caused a tsunami in the US

5d ago

1.1K

38



...



In Coding Beauty by Tari Ibaba

This new IDE just destroyed VS Code and Copilot without even...

Wow I never thought the day I stop using VS Code would come so soon...

Jan 17

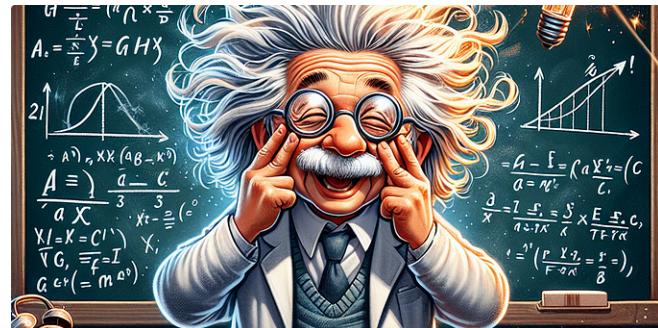
2.1K

85



...

Channel analytics



In Write A Catalyst by Awais Shahid

How Much YouTube Paid Me For 12 Million Views

How much you can expect from YouTube

Jan 14

7.1K

215



...



In The Knowledge of Laughter by David Graham

22 Hilariously Clever Jokes That Only the Most Intelligent Will Get

Think yourself intelligent, here's your chance to prove it

Jan 12

4K

103



...

[See more recommendations](#)