



Search models, datasets, users...



← Back to Articles

🔗 Open-R1: a fully open reproduction of DeepSeek-R1

Published January 28, 2025

[Update on GitHub](#)

▲ Upvote **689**



+683



[eliebak](#)

Elie Bakouch



[lvwerra](#)

Leandro von Werra



[lewtnun](#)

Lewis Tunstall

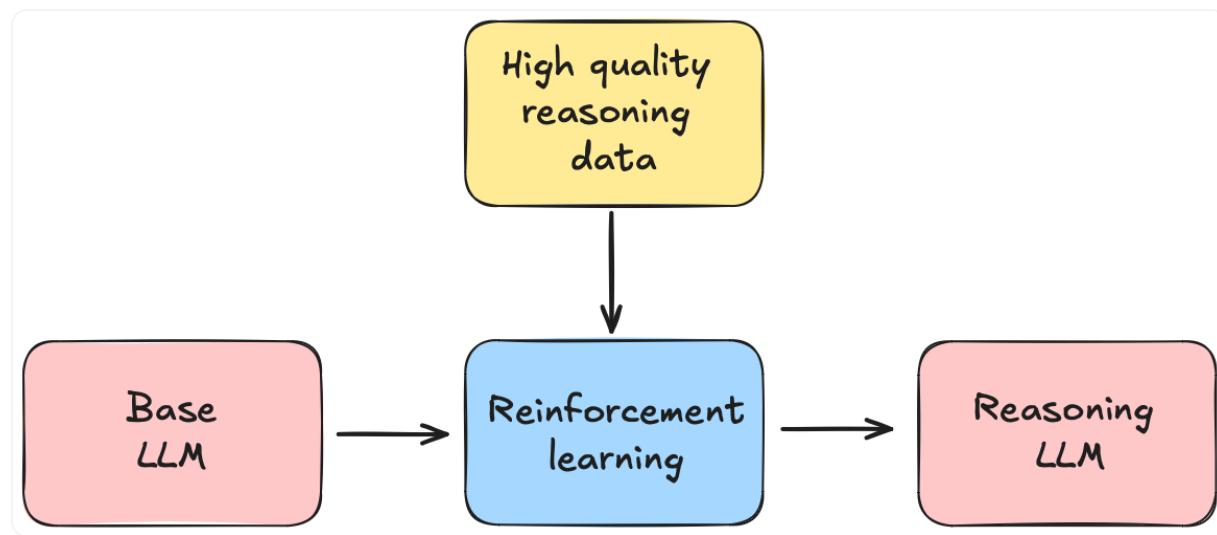
🔗 What is DeepSeek-R1?

If you've ever struggled with a tough math problem, you know how useful it is to think a little longer and work through it carefully. [OpenAI's o1 model](#) showed that when LLMs are trained to do the same—by using more compute during inference—they get significantly better at solving reasoning tasks like mathematics, coding, and logic.

However, the recipe behind OpenAI's reasoning models has been a well kept secret. That is, until last week, when DeepSeek released their [DeepSeek-R1](#) model and promptly broke the internet (and the [stock market!](#)).

Besides performing as well or better than o1, the [DeepSeek-R1](#) release was accompanied by a detailed [tech report](#) that outlined the key steps of their training recipe. This recipe involved several innovations, most notably the application of pure reinforcement learning to teach a base language model how to reason without *any* human supervision. As shown in the figure below, making a powerful reasoning

model is now very simple if you have access to a capable base model and a high-quality data mixture:



However, the DeepSeek-R1 release leaves open several questions about:

- **Data collection:** How were the reasoning-specific datasets curated?
- **Model training:** No training code was released by DeepSeek, so it is unknown which hyperparameters work best and how they differ across different model families and scales.
- **Scaling laws:** What are the compute and data trade-offs in training reasoning models?

These questions prompted us to launch the [Open-R1 project](#), an initiative to systematically reconstruct DeepSeek-R1's data and training pipeline, validate its claims, and push the boundaries of open reasoning models. By building Open-R1, we aim to provide transparency on how reinforcement learning can enhance reasoning, share reproducible insights with the open-source community, and create a foundation for future models to leverage these techniques.

In this blog post we take a look at key ingredients behind DeepSeek-R1, which parts we plan to replicate, and how to contribute to the Open-R1 project.

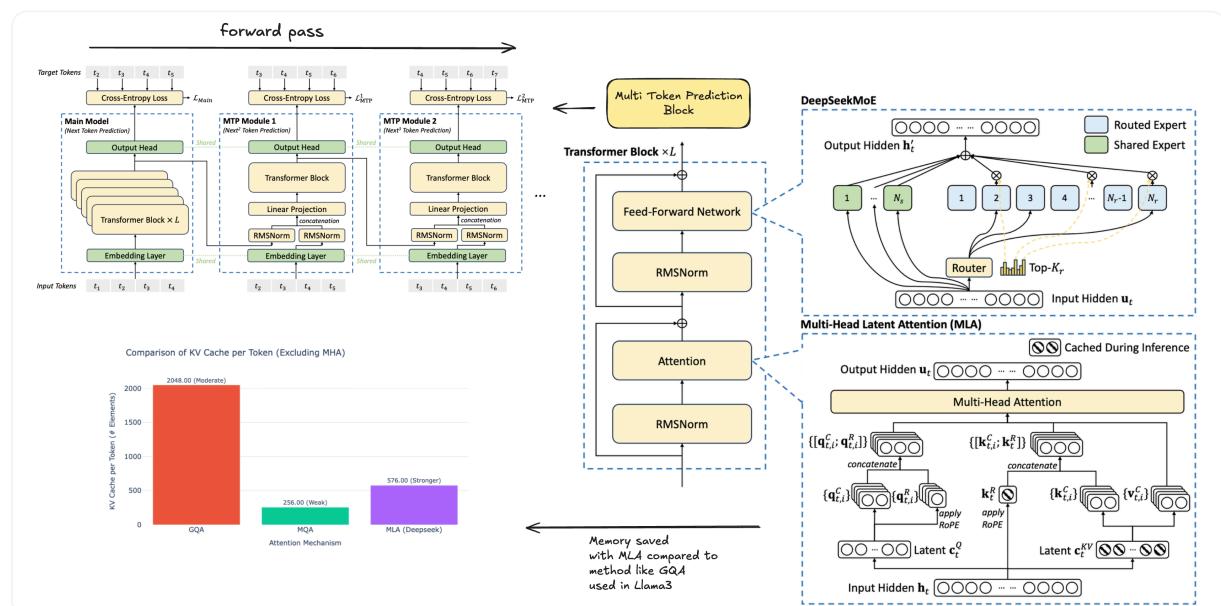
Let's dive in  !

🔗 How did they do it?

DeepSeek-R1 is a reasoning model built on the foundation of [DeepSeek-V3](#). Like any good reasoning model, it starts with a strong base model, and DeepSeek-V3 is exactly that. This 671B Mixture of Experts (MoE) model performs on par with heavyweights like Sonnet 3.5 and GPT-4o. What's especially impressive is how cost-efficient it was to train—just \$5.5M—thanks to architectural changes like Multi Token Prediction (MTP), Multi-Head Latent Attention (MLA) and a LOT (seriously, a lot) of hardware optimization.

DeepSeek also introduced two models: DeepSeek-R1-Zero and DeepSeek-R1, each with a distinct training approach. DeepSeek-R1-Zero skipped supervised fine-tuning altogether and relied entirely on reinforcement learning (RL), using Group Relative Policy Optimization (GRPO) to make the process more efficient. A simple reward system was used to guide the model, providing feedback based on the accuracy and structure of its answers. This approach helped the model develop useful reasoning skills, such as breaking problems into steps and verifying its own outputs. However, its responses often lacked clarity and were difficult to read.

That's where DeepSeek-R1 comes in. It started with a "cold start" phase, fine-tuning on a small set of carefully crafted examples to improve clarity and readability. From there, it went through more RL and refinement steps, including rejecting low-quality outputs with both human preference based and verifiable reward, to create a model that not only reasons well but also produces polished and consistent answers.



This all sounds great, but what's actually missing? Let's have a look at the missing pieces of the puzzle.

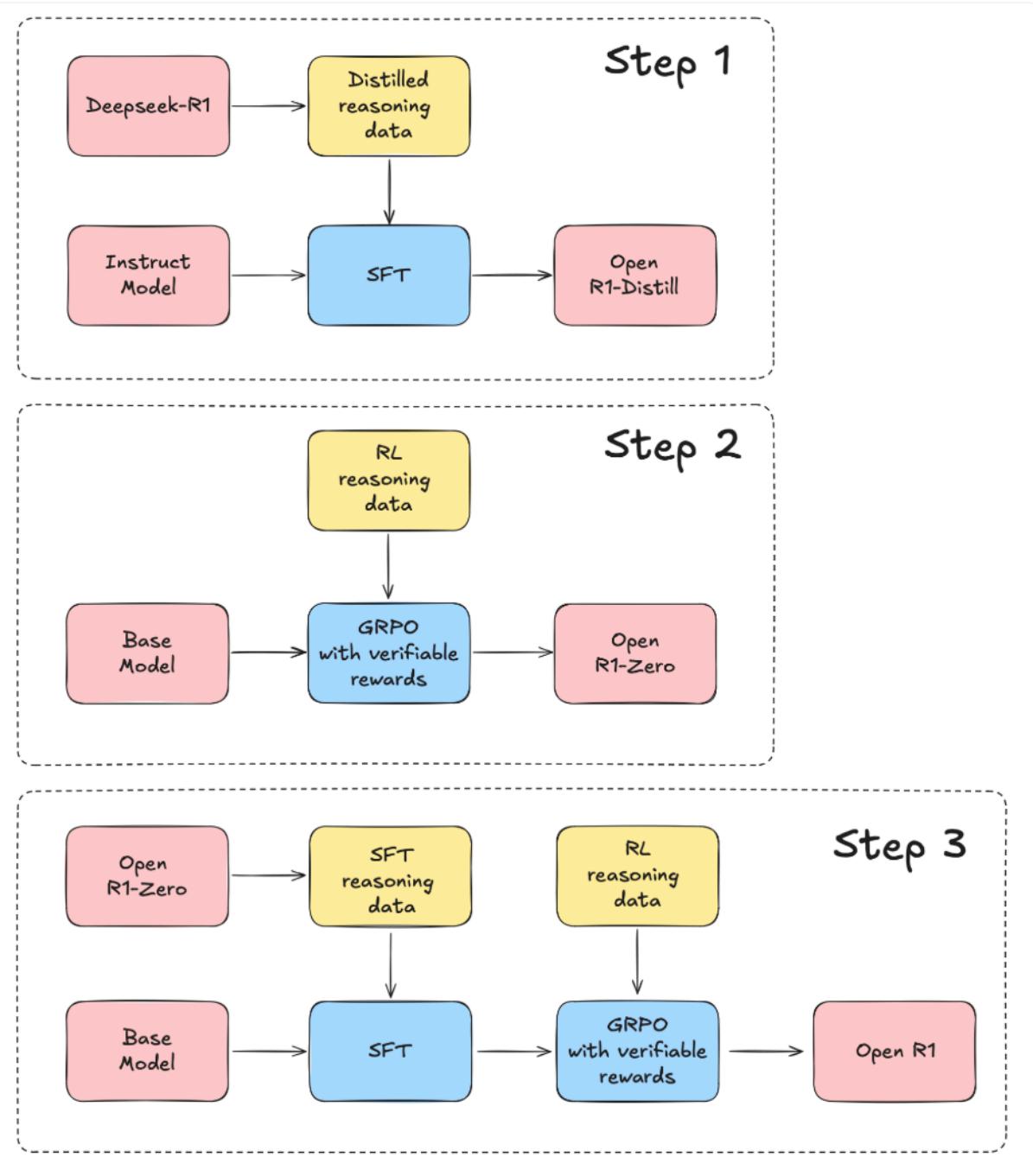
🔗 Open-R1: the missing pieces

The release of DeepSeek-R1 is an amazing boon for the community, but they didn't release *everything*—although the model weights are open, the datasets and code used to train the model are not 😰.

The goal of *Open-R1* is to build these last missing pieces so that the whole research and industry community can build similar or better models using these recipes and datasets. And by doing this in the open, everybody in the community can contribute!

As shown in the figure below, here's our plan of attack:

- **Step 1:** Replicate the R1-Distill models by distilling a high-quality reasoning dataset from DeepSeek-R1.
- **Step 2:** Replicate the pure RL pipeline that DeepSeek used to create R1-Zero. This will involve curating new, large-scale datasets for math, reasoning, and code.
- **Step 3:** Show we can go from base model → SFT → RL via multi-stage training.



The synthetic datasets will allow everybody to fine-tune existing or new LLMs into reasoning models by simply fine-tuning on them. The training recipes involving RL will serve as a starting point for anybody to build similar models from scratch and will allow researchers to build even more advanced methods on top.

Note that we don't want to stop at math datasets. There's a lot of potential in exploring other areas, obvious one like code but also scientific fields such as medicine, where reasoning models could have significant impact.

This initiative isn't just about replicating results—it's about sharing insights with the community. By documenting what works, what doesn't, and why, we hope to save

others from wasting time and compute on unproductive paths.

If this sounds interesting, we'd love your help! Whether it's contributing [code](#), joining discussions on [Hugging Face](#), there are plenty of ways to get involved. Let's build this together! 

More Articles from our Blog



DABStep: Data Agent Benchmark for Multi-step Reasoning

By eggie5 F.. • △ 38



Open Deep-Research – Freeing our search agents

By m-ric F.. • △ 871

👉 Community

 **facti** 13 days ago

⋮

Where is the evaluation numbers? without it you can't call it reproduction.

 8 replies •

 20

+

 **elephant-bear** 13 days ago

⋮

True, but it seems like there's nothing to be evaluated as of right now. I assume the ultimate goal is to train a new reasoning model and then use the same evaluation metrics as o1 and the DeepSeek-R1.

 18

+

> [Expand 7 replies](#)

[Reply in thread](#)

 **ghabib** 13 days ago • edited 13 days ago

⋮

That's quite interesting, I was asking myself why the questions the author exposed here are not being asked by others ? I believe the work they have done is memorable but at the same time I wonder why they wouldn't put these missing pieces on if they are supposed to be fully open.

Why even without reproduction and comprehension of the innovation they could affect so much the market in this way ?

 5 replies

 9

+

 **eliebak** Article author 13 days ago

⋮

Hi! This blog post is an introduction to the project, not a claim that we've reproduced R1 yet. We will totally share the missing piece when we have them, you can expect the models and datasets to be upload in this [Hugging Face org](#) and the code to be in this [GitHub repo](#)

 14

+

> [Expand 4 replies](#)

[Reply in thread](#)

 **bojan2501** 13 days ago

⋮

Interesting read, and it is good that we see more effort into this direction: more optimization and less brute force.

Also wonder what tool did the author use for creating step diagram.

 2 replies

 4

+

 **eliebak** Article author 13 days ago

⋮

Excalidraw 

 7

+

> [Expand 1 reply](#)

Reply in thread

● beretis 13 days ago

I'm so glad that initiative like this already exist, I'm gonna try to contribute :)

2 replies

3

1

1

+

● eliebak Article author 13 days ago

looking forward to it! 🚀

+

> Expand 1 reply

Reply in thread

● GameOsaku 13 days ago

So racist articel

2 replies

+

● schilling3003 13 days ago

WTF are your talking about?

+

> Expand 1 reply

Reply in thread

● ryanmarten 13 days ago • edited 13 days ago

Awesome to have this open reproduction started!

For Step #1 check out <https://github.com/open-thoughts/open-thoughts>!

<https://x.com/ryanmart3n/status/1884284101265612856>

Let's do this thing!

• 1 reply ·  5  5 

 **lvwerra** Article author 13 days ago 

It's really cool to see how the whole open source community comes together!

 7



Reply in thread

 **wsfung2008** 13 days ago 

Does anyone know the actual training cost of r1? I can't find it in the paper or the announcement post. Is the 6M cost reported by media just the number taken from v3's training cost?

• 2 replies ·  1 

 **ghabib** 13 days ago 

Ops...



> [Expand 1 reply](#)

Reply in thread

 **fnia** 12 days ago 

Has anyone asked the DeepSeek team to publish their training data and code, or at least share them privately with an independent replication project like this? Have they rejected such a request?

A faithful replication depends on using the same dataset and hyperparameters. Otherwise, any major discrepancies with the published benchmarks would be hard to pin down—whether due to training data differences or the replication method itself.

• 1 reply · 

 **lvwerra** Article author 12 days ago 

Historically, they have never released code or datasets of their LLM training, so I wouldn't expect this time to be different. If they would release it that would be amazing

of course!

In the meantime we have to make best guess estimates and see if we can get there ourselves.

+

[Reply in thread](#)

 **legolasyiu** 12 days ago

⋮

You provide good replication process of Deepseek reasoning training. I will try something similar to it.

 3

+

 Reply

 **nageshsomayajula** 12 days ago

⋮

This is really good information, can we fine tune with specific use case when code is released ?

 2 replies

 1

+

 **eliebak** Article author 12 days ago

⋮

Yes of course!

+

> [Expand 1 reply](#)

[Reply in thread](#)

 **appl044** 11 days ago

⋮

Please consider removing biased, tainted or unaligned training data and make an effort to remove copyrighted works from the crawl from intake. This will make the model more usable. If you reused anthropic curation checks, this might also help, remove obviouslybiased data will likely add a lot of value. We don't want another tainted, unaligned open source model, right? And no corporate would ever use deepseek or a model that reuses it, right?

We appreciate your work for the benefit of humanity, we hope.

Miike C from NJ

2 replies

+

enzomich 11 days ago

⋮

So basically you're asking to replace existing censorship with another flavour of censorship?



1

+

> Expand 1 reply

Reply in thread

PizzaHamburgerGuy 11 days ago

⋮

Can't wait! Hopefully the model will be uncensored but whatever you can do is alright! Love seeing open source building itself up. I'm not smart enough to actually help but I can contribute moral support lol



1

+

Reply

mseeger 10 days ago

⋮

Hello guys, I am even just trying to find code for DeepSeek-V2, in order to fully understand multi-head latent attention. You do not seem to have code in Hugging Face even for that. Or am I missing something? Don't see anything in src/transformers/models. MLA is not properly described in their paper, so it would be important to have code for this.

3 replies

+

lvwerra Article author 10 days ago

⋮

The code for the models are inside the model repositories, e.g. for V3:

https://huggingface.co/deepseek-ai/DeepSeek-V3/blob/main/modeling_deepseek.py



1

+

> Expand 2 replies

Reply in thread

pytholic 8 days ago

⋮

Is it possible to contribute to this project?

1 reply · 1 · +

 **eliebak** Article author 7 days ago

Yes, you can look at <https://huggingface.co/open-r1> and <https://github.com/huggingface/open-r1/discussions> :)

+

Reply in thread

 **ippali** 7 days ago

I wonder what data was used for the training of R1 and R1-zero. Is there any news on the possibly tos -breaking openAI API calls that were made this past autumn?

+

Reply

 **BILOO237** 7 days ago

Let's build this together

 2

+

Reply

 **charlieluо1999** 7 days ago

This project inspires me. At the moment, I developed an enhanced transformer. The goal is to figure out an equivalent SFT component in deepseek. The theoretical part and numeric examples worked out fine. I am now implementing it in a small scale LLM. Hopefully we can see the benchmark results in a few weeks. Thanks for the inspiration.

Charles

 2

+

Reply

 **moefear85** 6 days ago

releasing weights means open-binary ie freeware (as opposed to SAAS). but the actual blueprint/design of the network, defined in a HL language, is not provided, so it isn't even partially open (SOURCE).

+

Reply

Edit

Preview

Start discussing this article

 Tap or paste here to upload images

 Comment

 System theme

Company

TOS

Privacy

About

Jobs

Website

Models

Datasets

Spaces

Pricing

Docs

