

DeepSeek R1: An In-Depth Report

DeepSeek R1 represents a significant advancement in the field of reinforcement learning (RL) driven AI models. Developed by the Chinese AI firm DeepSeek, this family of models challenges established players like OpenAI and Google, offering a compelling combination of advanced reasoning capabilities and open-source accessibility ([GeeksforGeeks, 2025](#)). This report delves into the architecture, training process, performance benchmarks, and overall impact of DeepSeek R1, providing a comprehensive overview of its capabilities and limitations.

Architecture and Training

DeepSeek R1 builds upon the Mixture of Experts (MoE) architecture of its base model, DeepSeek-V3, utilizing a sparse activation approach that allows it to handle sequences up to 128,000 tokens ([UnfoldAI, n.d.](#)). This efficiency is evident in the model's ability to generate thousands of reasoning tokens per response while maintaining coherence and accuracy.

The core innovation of DeepSeek R1 lies in its unique training approach. Unlike many large language models (LLMs) that rely heavily on supervised fine-tuning (SFT), the initial model, DeepSeek-R1-Zero, employs pure reinforcement learning to develop reasoning capabilities ([UnfoldAI, n.d.](#)). This process begins with the base model and utilizes Group Relative Policy Optimization (GRPO), eliminating the need for a separate critic model. The GRPO implementation uses a reward function that balances accuracy and format adherence ([UnfoldAI, n.d.](#)).

DeepSeek R1's training pipeline consists of four distinct phases ([AI Papers Academy, n.d.](#); [Vellum AI, n.d.](#)):

1. **Cold Start:** The DeepSeek-V3-Base model is fine-tuned on a small, high-quality dataset of results generated by DeepSeek-R1-Zero. This addresses initial readability issues and provides a solid foundation for subsequent training. This dataset, while

containing thousands of samples, is considered relatively small in the context of LLM training ([AI Papers Academy, n.d.](#)).

- 2. Reasoning Reinforcement Learning:** This phase mirrors the training of R1-Zero, applying large-scale reinforcement learning to enhance reasoning skills, particularly in STEM fields, coding, and logic-based tasks ([AI Papers Academy, n.d.](#); [Vellum AI, n.d.](#)).
- 3. Rejection Sampling:** As RL approaches convergence, the model generates its own synthetic labeled data through rejection sampling. This involves selecting the best examples from the previous RL run, a technique reminiscent of strategies reportedly used by OpenAI ([Vellum AI, n.d.](#)).
- 4. Merged Supervised Fine-tuning:** The synthetic data generated in the previous step is combined with supervised data from DeepSeek-V3-Base in domains like writing and factual question answering. This final fine-tuning stage refines the model's performance across a broader range of tasks ([Vellum AI, n.d.](#)).

This multi-stage training process, incorporating both pure RL and SFT, distinguishes DeepSeek R1 from models relying solely on one approach. It aims to combine the raw reasoning power of RL with the refinement and polish achieved through supervised learning and human feedback ([Nguyen, 2025](#)).

Performance and Benchmarks

DeepSeek R1 has been evaluated against leading models like OpenAI's GPT series and Claude, demonstrating competitive performance across various benchmarks ([PromptHub, n.d.](#)). In reasoning and math tasks, R1 rivals or surpasses other models in accuracy and depth of reasoning. While GPT models generally perform better in coding benchmarks, R1 often outperforms in structured question-answering tasks. Notably, R1 excels in creative and long-context tasks, outperforming other models in benchmarks like AlpacaEval 2.0 and ArenaHard ([PromptHub, n.d.](#)). One key observation is that longer

reasoning chains generally improve performance, aligning with findings from other research ([PromptHub, n.d.](#)).

Strengths and Weaknesses

DeepSeek R1's strengths lie in its efficient architecture, innovative training methodology, and strong performance in reasoning tasks. Its open-source nature promotes accessibility and fosters community-driven development ([OpenTools AI, n.d.](#)). The model's cost-effectiveness is also a significant advantage, making advanced AI capabilities more accessible to a wider audience ([OpenTools AI, n.d.](#)).

However, R1 is not without limitations. Early versions exhibited issues with language mixing and less polished responses compared to chat-optimized models ([PromptHub, n.d.](#)). While these issues have been addressed through subsequent refinements, including supervised fine-tuning and human feedback, they highlight the ongoing development and iterative nature of LLM training. Concerns regarding potential biases in training data and the security implications of open-source models also warrant attention ([OpenTools AI, n.d.](#)).

Conclusion

DeepSeek R1 represents a compelling alternative in the rapidly evolving landscape of large language models. Its focus on reinforcement learning, coupled with a hybrid training approach, yields impressive reasoning capabilities. While challenges remain, the model's open-source nature, cost-effectiveness, and active development trajectory position it as a key player in the democratization of AI. Further research and community involvement will be crucial in realizing the full potential of DeepSeek R1 and addressing the broader implications of open-source AI models.

References

AI Papers Academy. (n.d.). *DeepSeek-R1 Paper Explained - A New RL LLMs Era in AI?*. Retrieved January 28, 2025, from <https://aipapersacademy.com/deepseek-r1/>

GeeksforGeeks. (2025). *DeepSeek Unveils DeepSeek-R1 RL Models: What's New and How It is better than OpenAI and Google*. Retrieved January 28, 2025, from <https://www.geeksforgeeks.org/deepseek-r1-rl-models-whats-new/>

Nguyen, T. N. (2025, January). *DeepSeek-R1: Architecture and training explain*. Medium. Retrieved January 28, 2025, from <https://medium.com/@namnguyenthe/deepseek-r1-architecture-and-training-explain-83319903a684>

OpenTools AI. (n.d.). *DeepSeek R1: The Open-Source AI Champion Giving OpenAI a Run for Its Money*. Retrieved January 28, 2025, from <https://opentools.ai/news/deepseek-r1-the-open-source-ai-champion-giving-openai-a-run-for-its-money>

PromptHub. (n.d.). *DeepSeek R-1 Model Overview and How it Ranks Against OpenAI's o1*. Retrieved January 28, 2025, from <https://www.prompthub.us/blog/deepseek-r-1-model-overview-and-how-it-ranks-against-openais-o1>

Tysoolen. (n.d.). *DeepSeek R1 vs OpenAI o1: The Ultimate Benchmark Comparison*. Retrieved January 28, 2025, from <https://www.tysoolen.com/story/deepseek-r1-openai-o1-ultimate-benchmark-showdown>

UnfoldAI. (n.d.). *DeepSeek-R1 — Training Language Models to reason through Reinforcement Learning*. Retrieved January 28, 2025, from <https://unfoldai.com/deepseek-r1/>

Vellum AI. (n.d.). *How DeepSeek-R1 Was Built; For dummies*. Retrieved January 28, 2025, from <https://www.vellum.ai/blog/the-training-of-deepseek-r1-and-ways-to-use-it>