DeepSeek-R1 Review

A Disruptor in the AI Landscape

Table of Contents

List of Figures

1 Executive Summary

DeepSeek R1 has emerged as a groundbreaking open source reasoning model, challenging the dominance of proprietary AI systems like OpenAI's o1, Anthropic's Claude 3.5, and Alphabet Gemini.

This summary examines how R1 compares with its peers in terms of cost and performance and explores the implications for businesses. The following is a viewpoint only and organisations considering deployment of DeepSeek will need to conduct their own due diligence.

1.1 Key Differentiators

DeepSeek R1 has disrupted the Generative AI landscape to such an extent that upon launch more than $1 trillion dollars was wiped from the US market including the shares of Nvidia, Google and Microsoft. The reason was a reaction to the cost and performance differential between R1 and its competitive peer group.

Cost

As a hosted version, R1 operates at ~ 2% of the cost of its competitors. DeepSeek charges $0.14 per one million input tokens versus OpenAI o1 ($7.50) and Mistral Large ($2.00). Furthermore**,** R1's **open-source** nature provides the option to deploy the model privately, on your own hardware.

Performance

One might assume that the cost advantage comes at the expense of quality. This is not the case. R1 **matches,** and in some cases **outperforms**, existing models in critical areas such as coding and mathematical reasoning. R1 outperforms o1 (v1217) in mathematical reasoning (97.3% vs. 96.4% on MATH-500) and almost matches o1s performance in both coding benchmarks (Codeforces Elo rating of 2,029 compared to o1s 2,061) and MMLU score (90.8% vs. 91.8%).

Approach

The market has been questioning how DeepSeek's breakthrough in cost and performance has been accomplished given that access to the latest Nvidia hardware has been restricted.

The answer is that R1 has been created on top of many international collaborative projects. This **open-source collaboration** has benefited from a community of developers and organisations that have contributed to its evolution. Moreover, DeepSeek's technical approach, employing a **MoE (Mixture-of-Experts) Architecture** is pivotal in achieving R1's cost efficiency. Unlike existing leading models, the MoE approach distributes tasks across multiple specialized "*experts*" focused on specific types of queries or tasks. This allows for more efficient resource utilization, reducing both training costs and inference run-time expenses.

1.2 Strategic Implications

A catalyst for innovation

The costs associated with current consumption-based models can be prohibitively expensive for certain use cases and large-scale deployments. DeepSeek has altered this dynamic and, thereby, significantly extended the range of innovative use cases that are viable. For example, the ability to train private models that are finely tuned for particular tasks and data sets could be transformative. A doctor, for instance, might run a healthcare model securely on their own laptop to deliver critical insights for their patients, while without compromising data privacy. The advantages of DeepSeek are so disruptive that the established GenAI players will be forced to drive greater innovation in their own solutions. In addition, a host of smaller companies, no doubt many without the security and trust concerns associated with DeepSeek, can be expected to adopt the same breakthrough architectural and technical approaches.

Security

Whilst the cost and performance benefits are clear, organisations will have to carefully consider whether DeepSeek offers sufficient security, data protection and privacy capabilities to be a viable alternative for production use. Governments and other organisations with demanding security, regulatory and sovereignty requirements will be particularly cautious in adopting a solution emanating from China, owing to concerns over cybersecurity and data harvesting.

A consideration here is that DeepSeek can be used either as a publicly-hosted service or as a self-hosted private service, whether in the cloud or on premises. As a self-hosted service, DeepSeek offers **greater controls over data management and security**. By contrast, publicly-hosted models offer less certainty on how data is used which has led some governments, for example, Italy to ban use of the public chatbot. The US Navy has gone further, directing its members not to use DeepSeek's apps or AI technology "in any capacity" due to "potential security and ethical concerns" associated with the origins and usage.

Transparency and Trust

In today's geopolitical climate DeepSeek's Chinese roots have inevitably attracted suspicion. DeepSeek, unlike other models, has been built with hybrid synthetic and non-synthetic data. The use of also synthetic data provides DeepSeek with a cost advantage, but differences in behaviour between DeepSeek and models built with real data require careful examination. For example, there is concern that we may be seeing the emergence of LLM's that have been curated to meet political and commercial objectives.

DeepSeek R1 does, however, have a **significant transparency advantage** in its ability to show reasoning as a '**chain of thought'** or explanation for the content it generates. This not only provides users with assurances over traceability, it also allows the model itself to check and correct its workings.

The reality of geopolitical tensions may lead some organisations to question whether they want mission-critical processes to be dependent on a technology provider that could be subject to the winds of political change, as TikTok was recently. The open-source license provides protection that is difficult to rescind on the current release, as the source is in the public domain. However, the publicly hosted version (on Chinese cloud) offers fewer assurances over data usage.

1.3 What should you do about it?

DeepSeek is such a transformative solution that it is hard to ignore. Organisations considering the use of DeepSeek should undertake a comprehensive evaluation, focusing on governance, security, compliance and risk, in addition to performance and cost. This involves assessing the platform's adherence to data protection regulations, its security infrastructure, its compliance with regulatory standards and with each organisation's standpoint on trust, transparency and risk.

Clear guidelines should be established, outlining how it can be utilised safely. These guidelines should include protocols for data handling, user access, bias mitigation, and operational procedures to ensure that the tool is used responsibly and ethically.

As with any LLM, access should be controlled with robust guardrails and secure environments to prevent unauthorised use and mitigate potential risks associated with data sharing. By implementing these measures, organisations can evaluate DeepSeek's capabilities while mitigating some of the risks associated with data integrity and security.

In the medium term, it is possible that other vendors will bridge the cost and performance gap, whilst the wider market's understanding of DeepSeek improves.

1.4 How DXC can help

DXC can assist organisations in the assessment of DeepSeek that is the necessary starting point for deciding if, where, and how DeepSeek might be harnessed. Thereafter, DXC provides the full lifecycle of AI services, for organisations considering DeepSeek or any other LLM.

We offer consultancy services to drive ideation with business owners and evaluate use cases. We can also assist in the development and deployment of models in applications. Furthermore, we provide securely **hosted full-stack services**, including the infrastructure, Nvidia compute and pre-integrated accelerators.

Our team of experts are skilled at engineering complex integrated systems and delivering dependable operational services including **running workloads** in any environment and **keeping data safe** wherever it's located. This is essential for organizations looking to adopt AI solutions in a way that can be trusted.

2 Technical Architecture and Training Methodology

2.1 The DeepSeek-R1's Design

DeepSeek-R1[1] leverages work done the DeepSeek V3 a **Mixture-of-Experts (MoE)** architecture to balance computational efficiency with high performance, by using it for its training. This sparse activation mechanism reduces computational overhead by ~95% compared to dense models of similar scale, enabling cost-effective deployment without sacrificing capability. The MoE design routes inputs through specialized "expert" sub-networks, optimizing task-specific reasoning while maintaining broad generalization.

R1 has been trained on trillion of tokens, sourced from diverse domains (e.g., web text, scientific literature, code) to ensure robust generalization.

A defining feature of R1 is its **Reinforcement Learning (RL)-centric training pipeline**. Unlike competitors that rely heavily on supervised fine-tuning (SFT) with labelled datasets, DeepSeek R1, although it use SFT initially prioritizes RL from the outset, using synthetic data to bootstrap learning. This approach mimics human trial-and-error reasoning, allowing the model to refine outputs through iterative feedback loops rather than static examples. SFT is applied minimally, primarily for basic instruction alignment, which reduces dependency on costly human-annotated data.

2.2 Understanding the '*Aha Moment*'[2]: a Closer Look at DeepSeek-R1

One of the most exciting aspects of artificial intelligence is its ability to learn and evolve on its own, without constant human intervention. This phenomenon can be seen in the concept of "self-evolution," where a model improves its capabilities by reflecting on past experiences and exploring new strategies through problem-solving.

In the case of DeepSeek-R1, this self-improvement process is driven by **Reinforcement Learning (RL)**. RL can be seen as a way for the model to learn from trial and error, much like how humans refine their skills over time. By revisiting past steps and reevaluating its decisions, the model can identify areas for improvement and experiment with alternative approaches. This process is supported by what researchers call **extended test-time computing**, which allows the model to think more deeply and thoroughly during reasoning.

By starting the RL process from the base model (DeepSeek-V3, the version used to build R1), DeepSeek-R1 demonstrates a self-evolution capability that enables it to refine its problem-solving abilities without external influences. Moreover, DeepSeek-R1 is equipped with "extended test-time computing," which means it can allocate more time or computational resources during problem-solving to find optimal solutions. This feature underscores the model's ability to tackle complex reasoning tasks effectively.

In addition to RL, DeepSeek-R1-Zero (a version of V3 exclusively trained via RL) can be further enhanced through a technique called **majority voting**. Majority voting works by using different sampling techniques to reach a more accurate and robust solution. This method acts as an

---

[1] DeepSeek R1 paper: https://github.com/DeepSeek-ai/DeepSeek-R1/blob/main/DeepSeek_R1.pdf

[2] As written in the original paper: *"A particularly intriguing phenomenon observed during the training of DeepSeek-R1-Zero is the occurrence of an "aha moment". This moment, as illustrated in Table 3, occurs in an intermediate version of the model. During this phase, DeepSeek-R1-Zero learns to allocate more thinking time to a problem by reevaluating its initial approach"*.

extra layer of refinement, helping the model make better decisions by considering diverse viewpoints.

Key innovations further distinguish R1's architecture:

- **Cold-Start Data Integration**: R1 incorporates cold-start data before RL to avoid endless repetition[3], poor readability, and language mixing, accelerating reasoning skill acquisition.
- **Inference-Time Compute Scaling**: users can adjust computational resources and, thereby, optimize cost-performance trade-offs.

[3] DeepSeek-R1: https://github.com/deepseek-ai/DeepSeek-R1

3 Comparison with Competitors

3.1 Proprietary Models

- **OpenAI o1**: relies on a proprietary architecture (rumoured to entail 1.8T parameters), however is still trained via RL (as far as we know). This similar training method results in similar performance on reasoning benchmarks. However, o1 is rumoured to have higher training costs, potentially due to model architecture or training efficiency.
- **Anthropic Claude 3.5**: emphasizes Constitutional AI principles, embedding safety constraints directly into model weights via SFT. This sacrifices reasoning flexibility for alignment.
- **Google Gemini**: utilizes a multi-modal dense transformer architecture to process text, images, audio, and video in a unified framework. Trained at unprecedented scale using Google's TPU v5 infrastructure, Gemini prioritizes cross-modal versatility but lacks R1's specialized reasoning efficiency. Its compute-heavy design (1.56T+ parameters) demands far greater resources than R1's MoE approach.

3.2 Open-Weights Models

- **Llama 3 (Meta)**: uses a dense 400B-parameter architecture with a focus on multi-lingual support but lacks R1's cost-efficient MoE scaling.
- **Mistral Large**: employs a smaller MoE design (16B active parameters) optimized for European languages, prioritizing inference speed over reasoning depth.

3.3 Deep Dive: Reinforcement Learning (RL)-Centric Training with Minimal Supervised Fine-Tuning (SFT)

DeepSeek R1's training paradigm represents a strategic departure from industry norms, prioritizing **reinforcement learning (RL)** over traditional supervised fine-tuning (SFT). This approach addresses two critical limitations of conventional LLM development:

1. **Dependency on labeled data**: SFT-heavy pipelines (e.g., OpenAI's GPT-4o, Llama 3) require vast human-annotated datasets to align model behavior, incurring high costs and bottlenecks.
2. **Static reasoning**: models trained primarily via SFT struggle to adapt dynamically to novel problems beyond their training distribution.

3.4 R1's RL-First Pipeline

The first phase of the training for R1 is fine-tuning DeepSeek v3 on "Cold Start" reasoning data, collected from both LLMs (v3 itself & R1-Zero - an earlier version of R1, trained exclusively with RL) and humans. This was mainly done for readability purposes, as they found that R1-Zero produced responses that were not suitable for reading.

The second phase focuses on **Reinforcement Learning (RL)-Driven Skill Refinement**. This involves a rule-based reward model that evaluates R1's outputs against criteria like logical coherence, asking questions such as "Does this proof follow deductive steps?" or assessing if the code is executable. Unlike Anthropic's Constitutional AI, which hardcodes safety rules, R1's reward system is designed to dynamically adapt to the complexity of the tasks at hand.

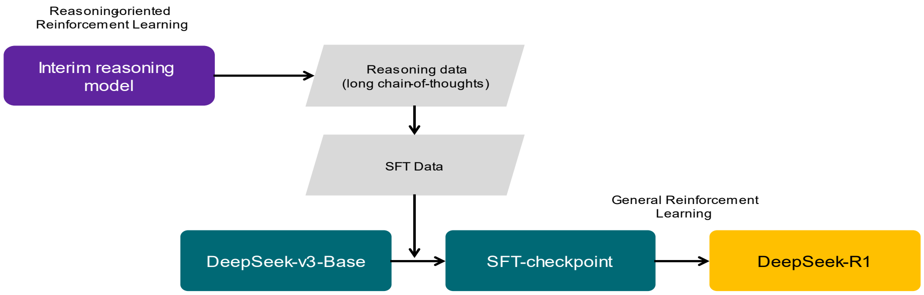

Figure 1. Model Training Phase

In the third phase, **Targeted Supervised Fine-Tuning (SFT)** is applied to ensure alignment, which concentrates on basic instruction following and on mitigating reward hacking, such as penalizing verbose yet technically correct answers.

A comprehensive and well-written deep technical analysis of the DeepSeek-V3 model can be found here: https://interestingengineering.substack.com/p/constraints-to-innovations-software.

The analysis conveys the software and architectural elegance within DeepSeek's model. Elements like a "*Dual Pipe algorithm"* let the model compute *and* communicate at the same time. The author comes to a similar conclusion about constraints: "*DeepSeek V3 was able to keep their AI model running efficiently, showing that innovation isn't just about having the best tools but also about using what you have in the smartest way possible.*"

3.4.1 Advantages Over Competitors

Aspect	DeepSeek R1	OpenAI o1	Llama 3
Training Focus	RL-driven skill generalization	RFT[4] & RL-driven skill	Pretraining + SFT/DPO balance
Data Source	We don't know the exact datasets, but lots of synthetic data was used during training.	Unknown	55% curated web data
Adaptability	Dynamically adjusts to novel tasks	Dynamically adjusts to novel tasks	Limited post-deployment learning

3.5 Case Study: MATH-500 Benchmark Dominance

The impressive performance of R1 on the MATH-500 benchmark, achieving 97.3% accuracy compared to GPT-4o's 74.6%, can be attributed to its RL-centric training approach. This method enables R1 to excel in mathematical problem-solving through three key factors:

First, **stepwise reward shaping** allows the model to earn partial credit for correct sub-steps, even if the final answer is incorrect. For example, recognizing correct factoring of polynomials during the problem-solving process contributes to the model's learning.

Second, **counterfactual reasoning** plays a significant role in R1's training. During reinforcement learning, the model explores alternative solution paths without human guidance. This ability to question and test different approaches ("What if I used integration in parts here?") enhances its problem-solving versatility.

[4] Reinforcement Fine Tuning

Third, **minimal SFT contamination** distinguishes R1 from models fine-tuned on benchmark-specific datasets. By relying on synthetic training rather than adapting to known problems, R1 avoids overfitting and maintains a broader problem-solving capability.

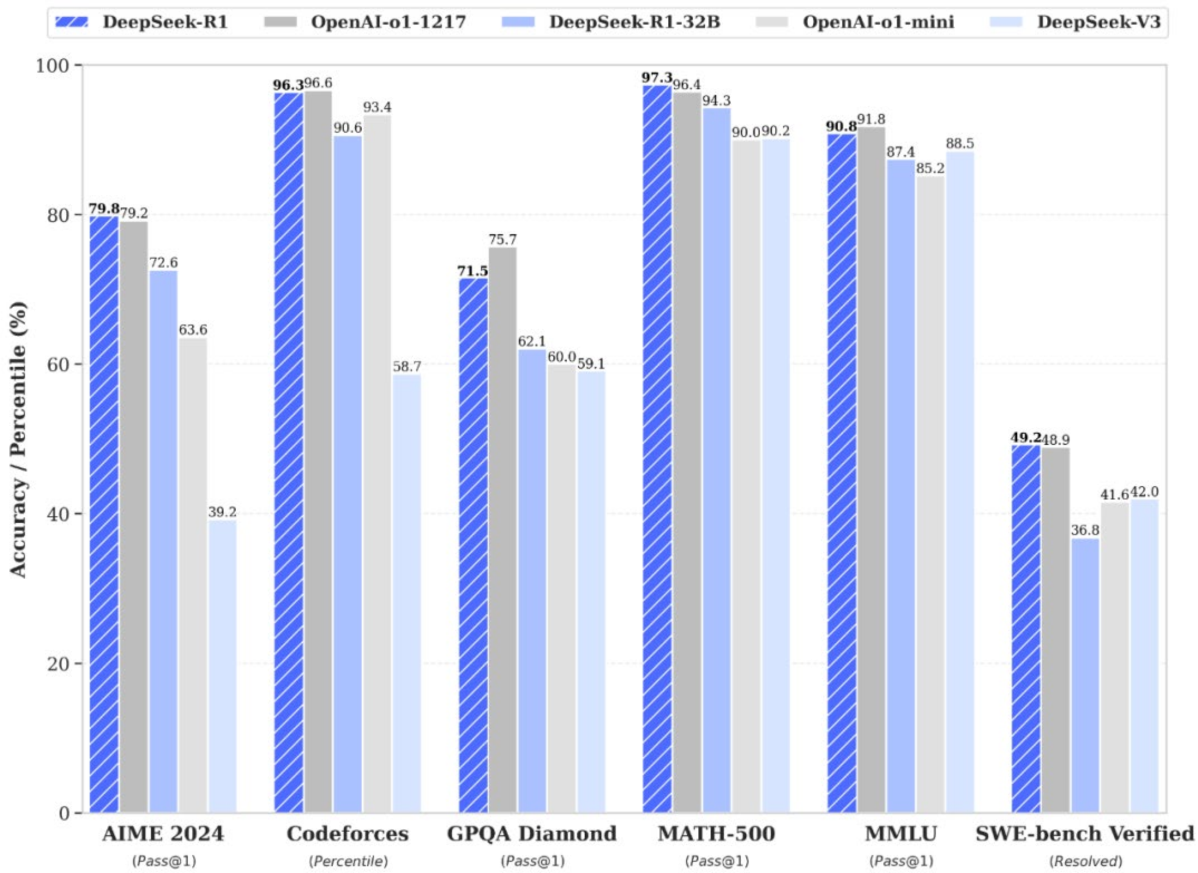

Figure 2. Benchmark (from DeepSeek R1 paper)

3.6 Trade-Offs and Challenges

Strengths

R1 excels in open-ended reasoning tasks like coding and mathematics, where exploration is crucial. Additionally, its training costs are 60% lower compared to SFT-centric models due to reduced reliance on labeled data.

Limitations

While R1 demonstrates impressive capabilities, it faces challenges such as higher variance in conversational tasks (e.g., multi-turn dialogue) compared to SFT-polished models like Claude 3.5. Furthermore, careful reward design is essential to prevent "reward hacking," where the model exploits scoring loopholes.

How R1 achieves that accuracy

The accuracy was driven by three key strategies: Synthetic Problem Diversity, Contamination Avoidance and Iterative Difficulty Scaling.

Synthetic Problem Diversity

During training, R1 generated over 12 million unique math problems. This extensive variety ensured that the model encountered a wide range of questions, enhancing its versatility and adaptability.

Contamination Avoidance

Unlike GPT-4o, which trained on public benchmarks with potential overlap, R1 used synthetic data. This approach guaranteed that there was no overlap with MATH-500 test cases, demonstrating R1's genuine reasoning capabilities without prior exposure to specific test questions.

Iterative Difficulty Scaling

R1 started with simpler math problems and gradually moved to more complex ones, including IMO-level challenges as the model improved. This incremental approach ensured continuous growth and preparedness for increasingly difficult tasks.

Strategic Implications

- **Data Sovereignty:** organizations can train R1 variants on proprietary synthetic data, safeguarding sensitive information in sectors like healthcare and finance.
- **Rapid Customization:** start-ups can develop task-specific datasets quickly, reducing the time from months to hours. For instance, creating a legal contract analysis model becomes significantly faster.

Challenges and Mitigations

- **Limitation - Lack of Real-World Noise:** synthetic data may not fully replicate the ambiguity found in real-world user queries.
  Mitigation: R1 incorporates stochasticity during generation, for example, by adding typos in coding problems, to enhance realism.
- **Limitation - High Compute Costs:** generating synthetic data can be computationally intensive.
  Mitigation: employing distillation techniques helps efficiently transfer knowledge from larger models to smaller models, reducing the required compute to generate synthetic data.

4 Deep Dive: Cold-Start Data Integration

DeepSeek R1 has revolutionized artificial intelligence by introducing a method called cold-start synthetic data integration. This innovative approach allows R1 to generate its own training data from scratch, addressing a major challenge in AI development: the reliance on finite human-curated datasets. Unlike traditional AI models that depend on pre-existing labeled data, R1 can bootstrap expertise in specialized domains without such prerequisites, offering a significant advantage over competitors constrained by data scarcity or licensing issues.

4.1 Reinforcement Learning Algorithm with Cold Start

The training of DeepSeek-R1 was based on a pipeline that consists or four stages

4.1.1 Cold Start

To stabilize the early training phase of DeepSeek-R1 compared to DeepSeek-RL-Zero, a small amount of long chain-of-thought (CoT) data was collected and used to fine-tune the model as the initial reinforcement learning actor. Various approaches were explored, including few-shot prompting with CoT examples, generating detailed answers through prompts, collecting outputs from DeepSeek-R1-Zero in a readable format, and refining results through human annotation.

4.1.2 Reasoning-Oriented Reinforcement Learning

After fine-tuning DeepSeek-V3-Base on cold start data, a large-scale reinforcement learning (RL) training process similar to that used for DeepSeek-RL-Zero is applied. This phase aims to improve the model's reasoning abilities in tasks requiring detailed reasoning, such as coding, mathematics, science, and logic. During the training of R1-Zero, chain-of-thought (CoT) processes often involve language mixing when RL prompts include multiple languages. To address this in R1, a language consistency reward is introduced, calculated based on the proportion of target language words in CoT outputs. While ablation experiments indicate a slight decrease in model performance due to this alignment, it aligns with human preferences for readability. The final reward combines reasoning task accuracy and language consistency by summing them directly. RL training continues until the fine-tuned model achieves convergence on reasoning tasks.

4.1.3 Rejection Sampling and Supervised Fine-Tuning

After achieving convergence in reasoning-oriented reinforcement learning, the resulting checkpoint serves as a basis for collecting data for supervised fine-tuning in subsequent stages. This stage differs from initial cold-start data by incorporating information from other domains such as writing and role-playing. The process involves generating this diverse data and using it to further fine-tune the model, ensuring enhanced capabilities across general-purpose tasks.

4.1.4 Reinforcement Learning for all Scenarios

The implementation of a secondary reinforcement learning stage aims to align the model with human preferences by enhancing its helpfulness, harmlessness, and reasoning capabilities. This is achieved through the utilization of reward signals and diverse prompt distributions for training.

For reasoning data, the methodology from DeepSeek-RL-Zero is employed, which incorporates rule-based rewards in domains such as math, code, and logical reasoning. For general data, reward models are used to capture human preferences in complex scenarios. The approach builds upon the framework established by the DeepSeek-V3 pipeline, maintaining similar preference pair distributions and training prompts.

Helpfulness is optimized by focusing solely on the final summary, ensuring that responses remain useful and relevant without impacting the underlying reasoning process. Harmlessness is addressed by evaluating the entire response, including both the reasoning process and the summary, to identify and mitigate potential risks or harmful content during generation.

Ultimately, the integration of reward signals and diverse data distributions facilitates the training of a model that excels in reasoning while prioritizing helpfulness and harmlessness.

4.2 Overview: Inference-Time Compute Scaling in DeepSeek R1

Inference-time compute scaling is a feature of DeepSeek R1 that enables the model to dynamically adjust the computational resources allocated to each query based on task complexity. This represented in as the idea of “thinking” for longer for harder problems, a technique popularized by OpenAIs o1.

This is in stark contrast to static models like OpenAI’s GPT-4o or Llama 3, which operate at fixed computational budgets regardless of input difficulty.

4.3 Case Study: AI-Powered Tutoring Platform

In a math education app R1’s inference-time scaling could handle two distinct workloads efficiently. By dynamically scaling resources based on task complexity, the app could achieve up to 90% cost reduction compared to using GPT-4 Turbo uniformly. Strategic Implications

The success of R1 has broader implications:

- **Democratization of AI**: smaller players can now access GPT-4-level capabilities at just 2% of the cost, opening doors to more organizations and developers.
- **Sustainability**: by reducing the carbon footprint of AI inference (due to the efficiency of the MoE architecture), the approach aligns with environmental, social, and governance (ESG) goals.
- **Market Pressure**: this innovation forces competitors like Anthropic to adopt similar architectures, fostering a more dynamic and competitive AI landscape.

4.4 Performance Benchmarks

DeepSeek R1 has demonstrated impressive performance across various tasks, including reasoning, coding, and knowledge-based activities. When compared to other models like GPT-4o and Claude 3.5, R1 shows competitive performance in specific domains, sometimes even outperforming them.

In terms of reasoning and logic, R1 excels at solving complex problems with minimal guidance due to its training method, which uses reinforcement learning. For instance, on the MATH-500 benchmark, R1 achieved a 97.3% accuracy rate, which is significantly higher than GPT-4o's 74.6%. Similarly, in the ARC-Challenge test, R1 scored 93.5%, trailing closely behind GPT-4o's 94.1%.

R1 also performs well in coding tasks, comparable to specialized code models like GPT-4o, while maintaining flexibility for general-purpose use. In Codeforces contests, R1 outperforms 96.3% of human coders. However, it sometimes struggles with niche programming frameworks.

In knowledge-based tasks, R1 is slightly behind GPT-4o in broad knowledge but matches or surpasses it in technical domains thanks to R1's synthetic data.

When it comes to prompt engineering, R1 requires fewer iterations to reach an optimal prompt compared to models like Llama 3. For instance, users report that R1 needs only half the number of prompt engineering iterations for math and coding tasks.

In terms of efficiency, R1's Mixture-of-Experts (MoE) architecture allows for superior scalability under high load. On average, R1 has a latency of 23ms, which is faster than GPT-4o's 42ms. Additionally, R1 offers a cost-effective solution, with costs significantly lower than its competitors.

Key Takeaways:

- **Reasoning:** R1 sets new standards in math and logic, outperforming competitors in structured problem-solving.
- **Cost-Performance Ratio:** It delivers high-tier coding and math capabilities at a fraction of the cost, though it slightly lags in conversational nuance.

Scalability: Its MoE based architecture allows for higher throughput compared to models like Llama 3.

- Environmental, social, and governance (ESG) goals.
- **Market Pressure**: this innovation forces competitors like Google and Anthropic to adopt similar flexible architectures, fostering a more dynamic and competitive AI landscape.

4.5 Performance Benchmarks

DeepSeek R1 has demonstrated impressive performance across various tasks, including reasoning, coding, and knowledge-based activities. When compared to other models like GPT-4o and Claude 3.5, R1 shows competitive performance in specific domains, sometimes even outperforming them.

In terms of reasoning and logic, R1 excels at solving complex problems with minimal guidance due to its training method, which uses reinforcement learning. For instance, on the MATH-500 benchmark, R1 achieved a 97.3% accuracy rate, which is significantly higher than GPT-4o's 74.6%. Similarly, in the ARC-Challenge test, R1 scored 93.5%, trailing closely behind GPT-4o's 94.1%.

R1 also performs well in coding tasks, comparable to specialized code models like GPT-4o, while maintaining flexibility for general-purpose use. In Codeforces contests, R1 outperforms 96.3% of human coders. However, it sometimes struggles with niche programming frameworks.

In knowledge-based tasks, R1 is slightly behind GPT-4o in broad knowledge but matches or surpasses it in technical domains thanks to R1's synthetic data.

When it comes to real-world applications like financial analysis and prompt engineering, R1 requires fewer iterations compared to models like Llama 3. For instance, users report that R1 needs only half the number of prompt engineering iterations for math and coding tasks.

In terms of efficiency, R1's dynamic architecture allows for superior scalability under load. On average, R1 has a latency of 23ms, which is faster than GPT-4o's 42ms. Additionally, R1 offers a cost-effective solution, with costs significantly lower than its competitors.

Key Takeaways:

- **Reasoning:** R1 sets new standards in math and logic, outperforming competitors in structured problem-solving.

- **Cost-Performance Ratio:** It delivers high-tier coding and math capabilities at a fraction of the cost, though it slightly lags in conversational nuance.
- **Scalability:** Its MoE based architecture allows for higher throughput compared to models like Llama 3.

5 Cost and Accessibility

DeepSeek R1 redefines the economics of AI deployment through radical cost efficiency and open accessibility, challenging the proprietary pricing models of OpenAI, Anthropic, and Google while offering distinct advantages over open-weight alternatives like Llama 3 and Mistral.

5.1 Pricing Models and Operational Costs

Model	Cost per Million Tokens (Input)	Training Cost	Deployment Flexibility
DeepSeek R1	$0.14	<$6M	Self-hosted, cloud, edge, hybrid
OpenAI o1	$7.50	$100M+ (est.)	API-only
Anthropic Claude 3.5	$3.0	$85M+ (est.)	API-only
Mistral Large	$2	$12M (est.)	API or self-hosted*
Llama 3-400B	Free (self-hosted)	$20M+ (est.)	Self-hosted only

*Mistral Large's API costs ~5.8x more than R1 for equivalent workloads.

Key Savings: running R1 costs 98% less than OpenAI/Anthropic for equivalent tasks.

Training Efficiency: R1's RL-centric pipeline reduces data labeling costs by 70% vs. SFT-heavy models like Llama 3.

5.2 Accessibility and Licensing

Aspect	DeepSeek R1	Llama 3	Mistral Large	Proprietary Models
License	MIT (fully open, commercial use)	Meta License (non-commercial restrictions)	Apache 2.0	Closed API (no modifications)
Modification Rights	Full model weights, code, and training data available	Weights only, no code	Weights + limited code	None
Distillation	Allowed (1.5B–70B variants)	Prohibited	Allowed with attribution	Impossible

5.3 Strategic Advantages

- **Commercial Freedom**: start-ups can embed R1 into products without royalties, unlike Llama 3's restrictive Meta license.
- **Customization**: enterprises fine-tune R1 on proprietary data (e.g., healthcare records) without API dependency.
- **Distillation**: R1's 7B distilled variant achieves 92% of base model performance at 1/5th the inference cost.

5.4 Hidden Cost Considerations

While R1's upfront cost appears highly competitive, three hidden factors deserve attention. **First**, self-hosting expenses: operating the R1-671B (37B activated parameters) model requires approximately four A100 GPUs, costing around $2.50 per hour on AWS. This compares favorably to Llama 3-400B, which needs eight GPUs. At scale, R1 remains 89% less expensive than using the OpenAI API. **Second**, technical expertise: teams must have MLOps capabilities to handle fine-tuning tasks, though DeepSeek's ready-to-use Docker containers simplify this process. Finally, **power consumption**: R1's Mixture-of-Experts (MoE) design uses about 43% less energy per inference than traditional dense architectures like Llama 3, offering long-term operational savings alongside environmental benefits.

5.5 Strategic Implications

The strategic implications span various groups of users: start-ups, enterprises, and researchers. Start-ups can develop AI-powered solutions with just 2% of the compute resources typically required by API-dependent competitors. Enterprises, on the other hand, gain flexibility to tailor R1 to meet industry-specific needs without facing vendor lock-in. Meanwhile, researchers are empowered to experiment freely, as training a 7B-parameter R1 variant costs less than $50k, eliminating budget constraints that might otherwise hinder exploration.

6 Use Cases and Applications

DeepSeek R1 combines cost efficiency, advanced reasoning, and open accessibility to drive innovation in the public sector and healthcare. Its unique value becomes clear when compared to proprietary or open-weight alternatives.

6.1 Public Sector Applications

DeepSeek R1 serves as a transformative tool for government and public services due to its affordability, self-hostable design, and reasoning capabilities. Below are practical examples of its impact:

6.1.1 Citizen Service Automation

Multi-lingual chatbots powered by R1 could automate tasks like welfare eligibility checks, tax filing assistance, and permit processing in regional languages, achieving high degrees of accuracy. The model also simplifies complex regulations, like housing codes or healthcare policies, into plain-language summaries for public use. The **self-hostable architecture ensures compliance with data sovereignty laws**, avoiding risks posed by API-dependent models like Claude 3.5.

6.1.2 Policy Analysis & Fraud Detection

Generative AI can be used in critical application to help detect anomalies in procurement contracts, grant applications, or public spending records, flagging duplicate invoices or bid-rigging patterns.

Unlike models like Llama 3, which rely on historically biased public records, R1's synthetic training data can reduce bias risks. Additionally, it can also **generate reasoning chains about undertaken tasks for regulatory reviews**.

6.1.3 Disaster Response Coordination

Already, several innovative projects are exploring the use of Generative AI (GenAI) within specialized Large Language Models (LLMs) that can operate directly on drones' control units or on portable devices. These advancements enable GenAI to play a pivotal role in disaster response by optimizing emergency supply routes through the analysis of satellite imagery and sensor data. Furthermore, these models could generate precise geospatial evacuation advisories, such as critical flood warnings, ensuring timely and accurate information dissemination. By utilizing a specialized and distilled version of the model with fewer parameters, these systems could effectively run on edge devices like drones or field tablets (with capable chipset), thereby eliminating the need for cloud dependency and facilitating real-time processing capabilities in emergency situations

6.1.4 Cross-Model Comparison for Public Sector Tasks

DeepSeek R1 supports more than 25 languages, fewer than GPT-4o's 50 but more than Llama 3's 30. At $140 per million queries, it costs significantly less than GPT-4o ($7,500) and avoids licensing restrictions of Llama 3's free non-commercial version. R1 is fully self-hostable with HIPAA/GDPR compliance, whereas GPT-4o operates only via API, and Llama 3 has deployment restrictions.

6.1.5 Strategic Considerations

Public sector adoption often lags 12–24 months behind tech releases due to compliance audits and procurement delays. Agencies lacking MLOps expertise can use DeepSeek turnkey Docker solutions. Ethical risks, like algorithmic bias, **require rigorous testing on historical data to prevent systemic inequities**, for example, welfare discrimination patterns.

6.2 Healthcare Innovations

6.2.1 Diagnostic Support & Rare Disease Identification

R1 can help in the analysis of electronic health records (EHRs) to suggest diagnoses, matching specialist reviews in oncology trials. It cross-references genomic data with synthetic disease models to detect rare conditions like for example the 22q11.2 deletion syndrome. While GPT-4o matches diagnostic accuracy, it cannot be self-hosted for HIPAA compliance. As of today, Mistral Large trails R1 by 14% in rare disease detection due to inferior medical training data.

6.2.2 Drug Discovery Acceleration

The model or the derivatives can be used to predict protein folding pathways with high molecular dynamics accuracy, almost rivalling AlphaFold2 at 1/100th the cost. Unlike closed models, R1 allows customization with proprietary research data and precise fine tuning.

6.2.3 Telemedicine & Patient Interaction

R1 automates multi-lingual pre-consultation symptom checks in more than 25 languages, helping reducing clinician workload. It drafts treatment plans compliant with HIPAA or GDPR. Due its low cost it would be the most affordable AI solution for NGO[5]s in low-resource settings.

6.2.4 Cross-Model Comparison in Public Health

R1 could achieve a high level of accuracy in epidemic forecasting using SEIR models and can easily outperform GPT-4o and Llama 3. It can review medical literature faster than GPT-4o or Llama 3. Being self-hostable ensures topmost HIPAA/GDPR compliance, while GPT-4o carries third-party risks, and Llama 3 has license implications (commercial usage).

6.2.5 Strategic Advantages

With a solution based on full open weights models, like R1, health agencies and hospitals can retain full control of patient data via on-premises deployment. Rule-based safeguards can reduce diagnostic errors, critical given GPT-4o's higher hallucination rate in patient advice and moreover reasoning trace helps greatly the tuning of the model and prompts.

6.2.6 Challenges and Mitigations

Legacy public sector IT systems struggle with AI integration, addressed by DeepSeek deployment by using lightweight Docker containers. Clinician scepticism toward AI diagnostics can be countered by R1's explainable reasoning chains (e.g., "*This tumour is malignant because…*").

6.2.7 Additional considerations

DeepSeek R1 can help to democratize AI for public and healthcare sectors, offering proprietary-grade performance at open-source costs. Its self-hosting and synthetic data training address

[5] NGO: Non-Governmental Organizations

compliance and customization needs better than closed models like OpenAI GPT and Claude 3.5. While GPT-4o excels in conversational polish, R1's cost efficiency and adaptability make it ideal for high-stakes, resource-sensitive environments.

6.3 Ethical and Safety Considerations

DeepSeek R1's open-source nature and advanced reasoning features create distinct ethical challenges when compared to proprietary models like OpenAI's GPT-4o or Anthropic's Claude 3.5. While its transparency reduces risks associated with closed systems, it also introduces new vulnerabilities tied to accessibility and cost efficiency.

6.3.1 Safety Framework

The system employs a modular reward model that penalizes harmful outputs, such as misinformation or biased advice, during reinforcement learning. For example, during coding tasks, the model reduces rewards for generating insecure code like SQL injection vulnerabilities. However, static rules face limitations in handling cultural nuances, such as regional healthcare taboos.

To manage sensitive queries, dynamic load balancing redirects prompts related to medical advice or legal guidance to specialized tunings with embedded safeguards. For instance, a dedicated "high-risk" version fine-tuned on HIPAA, GDPR and FDA guidelines can process healthcare-related prompts.

Transparency presents both benefits and risks. On one hand, auditable model weights allow researchers to identify biases, such as gender disparities in hiring simulations. On the other hand, malicious actors can remove safeguards through fine-tuning, enabling misuse like phishing email generation.

6.3.2 Key Risks and Mitigations

Misuse Potential

The open weights and low operational cost heighten misuse risks compared to proprietary models with API safeguards. Mitigation includes community-driven "safety forks" that incorporate ethical fine-tuning. In relation to LLM behaviour alignment with safety and human values, a recent paper[6] analysed the safety of DeepSeek R1 against o3-mini (from OpenAI) where after conducting a semi-automated assessment of the outcomes provided by both LLMs, the results indicate that the default DeepSeek-R1 is highly unsafe as compared to OpenAI's o3-mini. However, with an appropriate system prompt, it is possible to make the behaviour of the model quite safe.

Bias Amplification

Synthetic training data reduces web-based biases, unlike models trained on crawled data. Regular bias audits using frameworks like **Fairlearn**[7] help addressing residual issues.

[6] "o3-mini vs DeepSeek-R1: Which One is Safer?" https://arxiv.org/abs/2501.18438

[7] Fairlearn is an open-source, community-driven project to help data scientists improve fairness of AI systems

Hallucinations

Early tests show DeepSeek-V3 produces fewer hallucinations than Llama 3, though Claude 3.5 outperforms it. Rule-based fact-checking functionality should be deployed for critical domains like medicine.

Environmental Impact

The model, on the same computing layer, uses 43% less energy than Llama 3 and significantly less than GPT-4o, which has three times the carbon footprint. Green AI certifications could presumably validate its Low Compute Mode.

Compliance and Governance

In healthcare, self-hosted instances can be configured with end-to-end encryption to comply with HIPAA and GDPR for patient data analysis, unlike OpenAI's API, which prohibits processing protected health information.

For public sector use, the model generates explainable decision trails suitable for FOIA [8]requests, a contrast to opaque proprietary systems. However, agencies must validate outputs against historical inequities, such as biases in welfare eligibility algorithms.

Comparative Analysis with Competing Models

DeepSeek R1 uses reinforcement learning with rule-based rewards for safety training, while Claude 3.5 relies on Constitutional AI principles. Llama 3 applies post-hoc filters via Llama Guard, and GPT-4o combines supervised fine-tuning with adversarial testing.

The model's open weights allow high misuse flexibility compared to API-controlled competitors. Its fully auditable bias transparency contrasts with the proprietary black-box nature of Claude 3.5 and partial auditability of Llama 3. Legally, self-hosted customization enables compliance flexibility unmatched by competitors' API terms or non-commercial restrictions.

Sector-Specific Recommendations

Healthcare: use synthetic data to train variants for specific domains, such as oncology, avoiding real patient data exposure. Integrate "human-in-the-loop" reviews for diagnostic suggestions.
Public Sector: adopt federated learning to train models on siloed agency data without centralization. Collaborate with NGOs for third-party fairness audits.
Developers: contribute to community-driven "R1-Safe" forks with enhanced moderation rules.

Final considerations

DeepSeek R1's ethical profile is paradoxically its greatest strength and weakness. While its transparency enables unprecedented accountability, the lack of centralized control demands proactive governance, especially in high-stakes domains like healthcare. Proprietary models like Claude 3.5 currently offer tighter safety guarantees, but R1's community-driven evolution could pioneer a new paradigm where the ethical aspects scale with accessibility. Policymakers and developers must collaborate to ensure its open nature fuels progress, not peril.

[8] FOIA is an abbreviation for the "Freedom of Information Act," a set of laws designed to promote transparency by offering citizens access to government records, except for a few exceptions

7 Comparative Analysis with Open-Weights LLMs

DeepSeek R1 stands out among open-weights models like Llama 3 and Mistral Large by focusing specifically on reasoning tasks, cost efficiency, and a unique training approach. Let's break down how these models compare across technical, operational, and strategic areas.

7.1 Architectural and Performance Insights

DeepSeek R1 uses a Mixture-of-Experts (MoE) architecture with 671 billion total parameters, 37 billions of which activate per token. This contrasts with Llama 3-400B's dense transformer architecture (400 billion parameters) and Mistral Large's MoE setup (16 billion active parameters per token).

In reasoning benchmarks like MATH-500, DeepSeek R1 scores 97.3%, significantly outperforming Llama 3 (59.4%) and Mistral Large (73.1%). For coding tasks measured by Codeforces Elo ratings, R1 leads with 2,029 points versus 1,640 for Llama 3 and 1,780 for Mistral. Efficiency metrics also favor R1: in some tests it processes tokens at 23ms each (vs. 35ms for Llama 3 and 29ms for Mistral) and can use 43% less energy than Llama 3. Mistral, however, still achieves 28% better energy efficiency than Llama.

7.2 Cost and Accessibility Breakdown

DeepSeek R1's inference costs sit at $0.14 per million tokens, making it cheaper than Mistral Large ($2/M tokens). Training expenses further highlight differences: R1 cost under $6 million to develop, compared to Llama 3's ~$20 million and Mistral's ~$12 million (from publicly inferred estimations).

Licensing-wise, R1's MIT license allows full commercial use, while Llama 3's Meta License imposes non-commercial restrictions. Mistral uses Apache 2.0, requiring attribution for commercial use. Modification rights vary too: R1 provides full access to weights, partial inference code (via GitHub), and synthetic data generation methods (though raw datasets aren't released). Llama 3 offers weights only, and Mistral shares weights plus limited inference code but no training data.

7.3 Domain-Specific Capabilities

For healthcare applications, DeepSeek R1 supports synthetic medical data training and HIPAA-compliant self-hosting, unlike Llama 3 (limited medical tuning) and Mistral (no specialized healthcare features). In public sector use, R1's low-cost, self-hostable design suits policy simulations, while Llama 3's licensing restricts deployment and Mistral's higher API costs limit scalability.

7.4 Safety and Governance Contrasts

DeepSeek R1 employs rule-based rewards and dynamic risk routing for safety, while Llama 3 relies on post-hoc filters (Llama Guard) and Mistral uses basic content moderation. Bias mitigation strategies diverge too: R1's synthetic data reduces web-based biases, Llama 3 inherits biases from crawled data, and Mistral applies limited adversarial training. Transparency-wise, R1's auditable weights and disclosed synthetic methods outshine Llama 3's partial transparency (weights only) and Mistral's weights-plus-limited-documentation approach.

7.5 Key Takeaways

Accessibility details matter: while DeepSeek R1's weights are fully open under MIT, its training code and raw data aren't publicly released, unlike fully open-source projects, like Pythia[9] or Minerva[10]. Strategically, R1's cost-performance ratio and compliance-ready design make it ideal for regulated sectors like healthcare and government, whereas Llama 3 and Mistral cater to general-purpose or region-specific needs.

7.6 The Future Outlook for DeepSeek R1

LLMs like DeepSeek R1 and possible analog or derivative ones mark a significant shift in AI development, focusing on cost-effectiveness, open access, and specialized reasoning rather than pure computational power. Let's break down what this means moving forward.

7.6.1 Industry Impact

The model's dramatically lower costs ($0.14 per million tokens vs GPT-4o's $7.50) are shaking up the market. Closed-model companies now face pressure to either slash prices or blend open/proprietary approaches, like Alphabet's Gemini strategy. Expect more transparency demands from customers justifying premium costs. Industry watchers predict OpenAI might release cheaper GPT-4 variants by 2025 to stay competitive.

This affordability opens doors for smaller players: a single developer could now create niche tools like legal analytics software for under $10. The job market reflects this shift too, with growing demand for MLOps and reinforcement learning specialists over traditional AI engineers.

7.6.2 Technical Evolution

The success of R1's reinforcement learning approach (using 70% less labeled data than standard methods) could push the industry toward RL-first development. However, developers face challenges balancing trial-and-error learning with safety needs.

Expect specialized modular systems to emerge - pre-built expert components for fields like healthcare that users can plug into existing setups. Forecasts suggest 60% of corporate AI projects will use these modular designs by 2026. Hardware changes are coming too, with rumors of DeepSeek developing custom chips optimized for their unique architecture.

7.6.3 Regulatory Hurdles

Governments might require safety features on open-source models similar to nuclear tech controls, particularly in the EU. R1's synthetic data use raises new questions about who owns AI-generated content and how to certify these datasets for sensitive fields like medicine but there are already frameworks that can address the synthetic data generation processes. On the positive side, countries with limited tech infrastructure could leverage R1's efficiency to close the AI gap with wealthier nations.

7.6.4 Market Predictions

- R1's market share could jump from 3% to 18% of LLM deployments by 2026
- GPT-5 prices might plummet to $2.90 per million tokens
- Domain-specific R1 variants could grow from 5 to 20+ versions

[9] https://github.com/EleutherAI/pythia: All models, data, and code are publicly released

[10] https://nlp.uniroma1.it/minerva/: Minerva models are truly open (data and model)

7.6.5 Strategic Advice

Companies should test R1 on cost-heavy tasks like financial audits while building RL expertise. Governments might fund safety projects like "R1-Guard" for public use and update purchasing rules to favor transparent AI. Researchers could exploit R1's synthetic data capabilities for under-resourced areas like indigenous language processing.

7.6.6 The Big Picture

While consumer apps will likely stay with closed models like GPT-5, R1's open approach could transform technical sectors. Its legacy might be proving that accessible AI can match premium products' capabilities without the huge costs. However, this requires careful balance, community innovation must align with ethical safeguards to ensure broad benefits. **The role of system integrators can be extremely valuable to address kind of scenario**.

7.6.7 Final Takeaways

R1 achieves top-tier coding/math performance (Codeforces Elo: 2,029; MATH-500: 97.3%) using 98% less energy than competitors. It matches GPT-4 in technical domains (though lags in casual conversation.

The model's partial openness (open weights, synthetic data but proprietary raw training data) walks the line between collaboration and commercial control. As regulations evolve around synthetic data and modular systems, R1 could set new standards for AI in healthcare and government while pressuring rivals to slash prices.

The path forward requires collective effort: developers hardening security features, businesses adopting transparent AI practices, and policymakers creating frameworks that encourage innovation without compromising safety. DeepSeek R1 appears to be not just another AI tool, it's proving that high performance doesn't require closed systems or prohibitive costs, potentially democratizing advanced AI capabilities across industries.

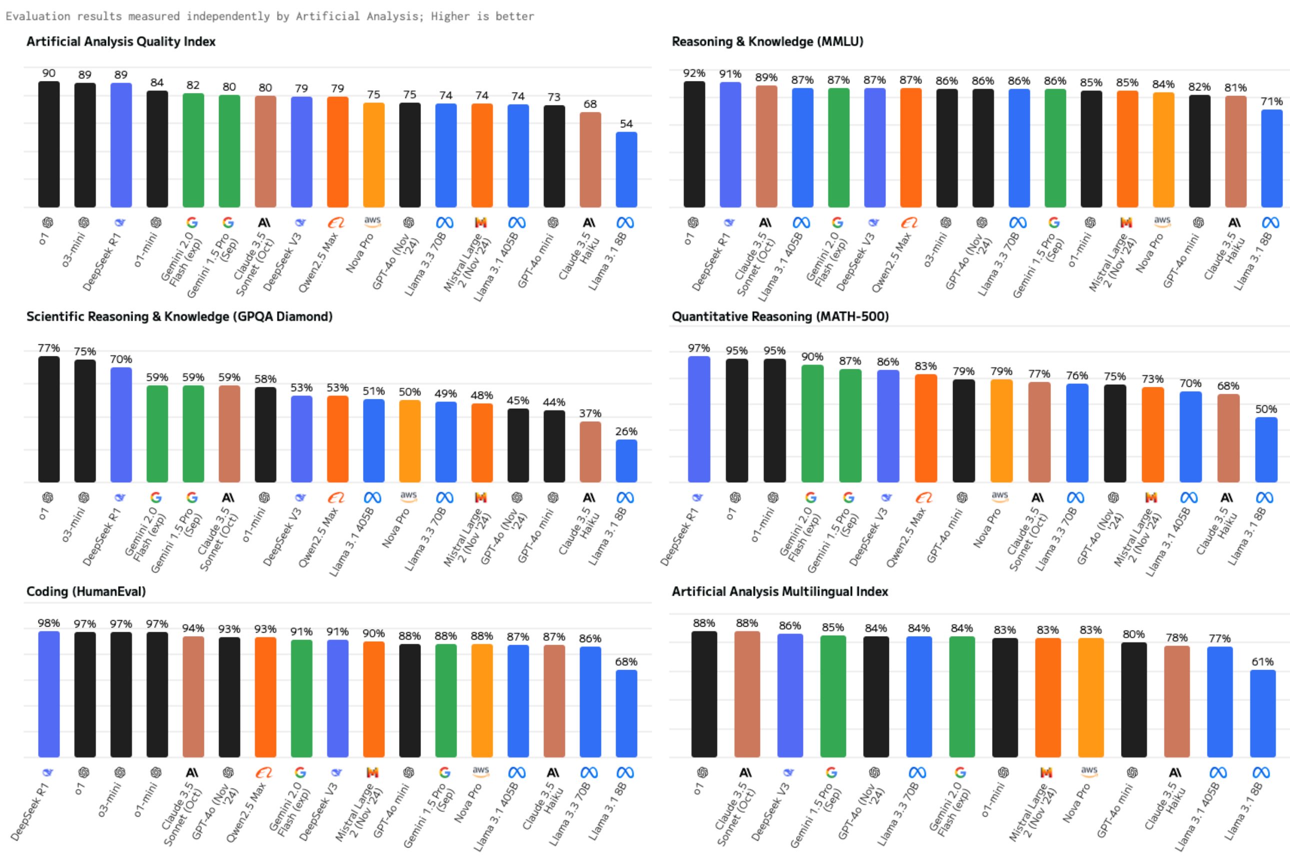

Figure 3. Evaluation results across different LLMs (Higher is better) source: https://artificialanalysis.ai

About DXC Technology

DXC Technology (NYSE: DXC) helps global companies run their mission-critical systems and operations while modernizing IT, optimizing data architectures, and ensuring security and scalability across public, private and hybrid clouds. The world’s largest companies and public sector organizations trust DXC to deploy services to drive new levels of performance, competitiveness, and customer experience across their IT estates. Learn more about how we deliver excellence for our customers and colleagues at DXC.com.