



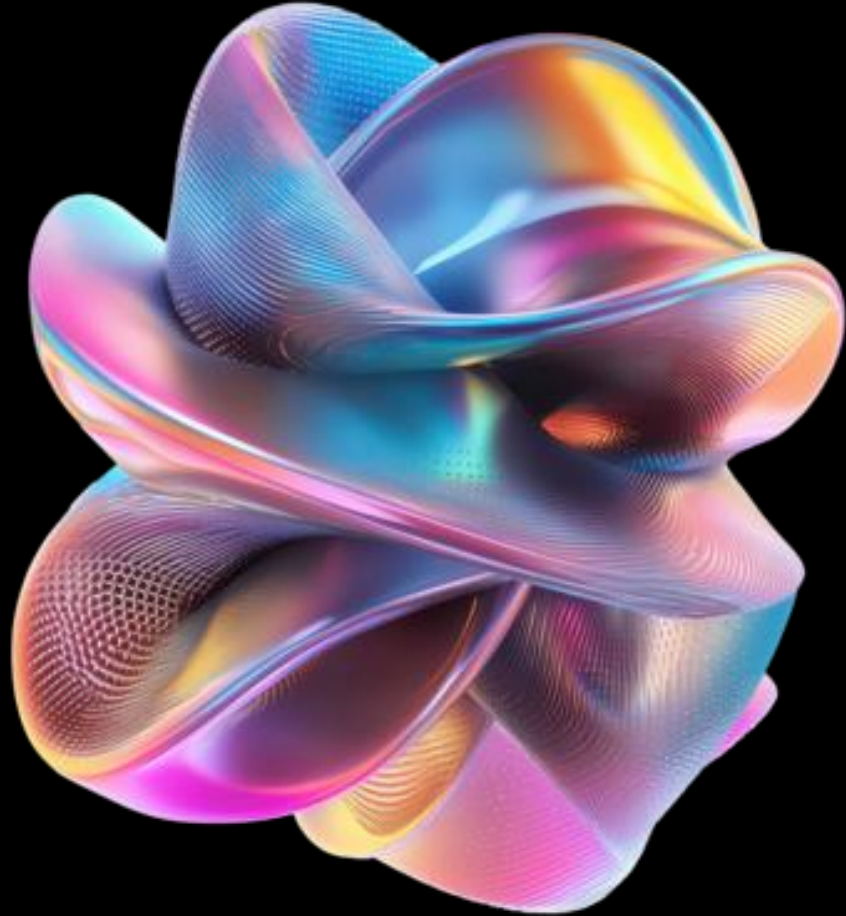
Microsoft Research

Phi-4 Reasoning

Power Through Precision

Mojan Javaheripi

Senior Researcher, Microsoft Research AI Frontiers



Reasoning Models

- What are reasoning models?
- Why are they important?



Core Characteristics of Reasoning Models

Models that follow explicit logical steps to solve problems.



Problem Decomposition

Breaking complex tasks into simpler sub-problems



Iterative Refinement

Self-correction capabilities



Explainable Processes

Transparent reasoning paths



Verifiable Outputs

Solutions that can be checked step-by-step

Reasoning Models are designed for Complex Decisions

Mathematical Reasoning

Solving equations and proofs



Algorithmic Thinking

Developing systematic approaches



Strategic Planning

Multi-step problem solving



Logical Analysis

Evaluating complex arguments





The Phi Reasoning models

Phi-4 Reasoning Models



Base Model: Phi-4

14B parameter dense decoder-only Transformer architecture



Phi-4-Reasoning

Fine-tuned on 1.4M+ STEM and coding prompts



Phi-4-Reasoning-Plus

Further enhanced with GRPO reinforcement learning

David vs Goliath: Size vs Intelligence

5-50x

Size Difference

Competing with giants

Competes with models many times larger while maintaining superior efficiency and performance per parameter.

32K

Context Length

Extended reasoning window

Extended token window enables complex reasoning chains and comprehensive problem analysis.

14B

Parameters

Efficient architecture

Compact model size delivering outsized performance through strategic training and optimization.

Strong Community Adoption

Phi-4-reasoning

30K

Downloads on
Hugging Face

84 quantized models

Phi-4-reasoning-plus

24K

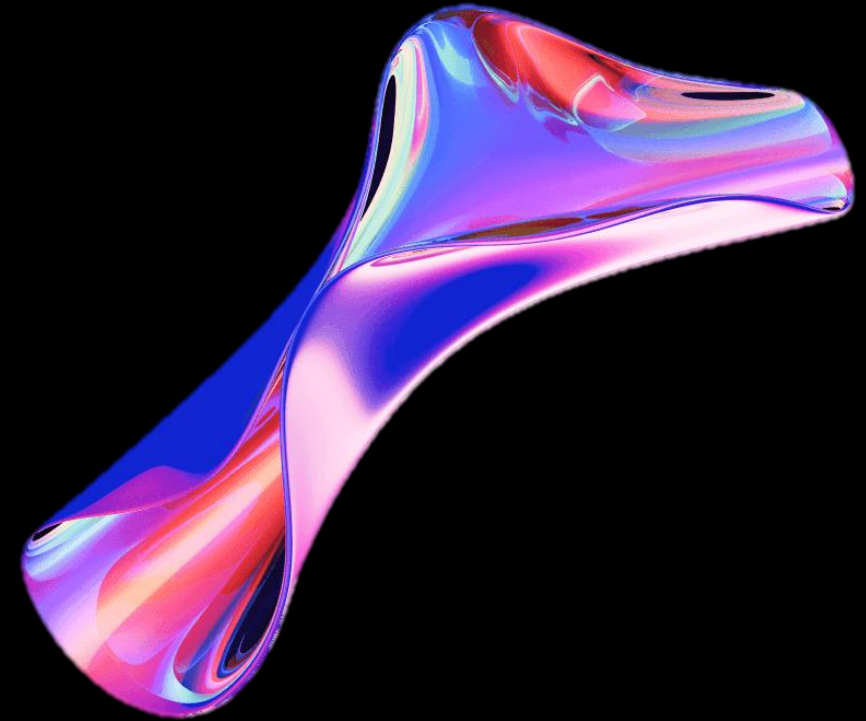
Downloads on
Hugging Face

26 quantized models







Unsloth GGUF
Version

120K

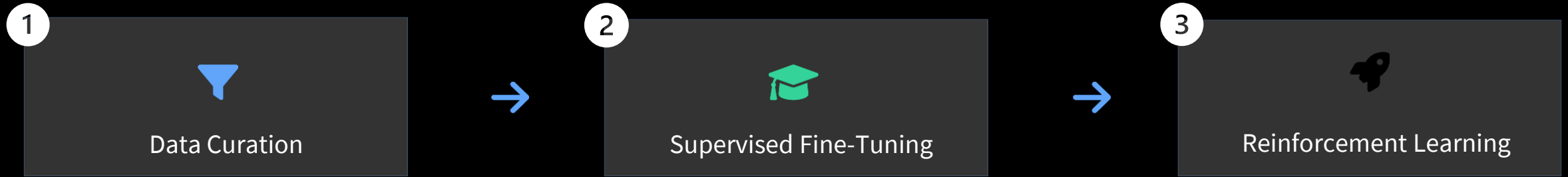
Most popular
quantized variant



Strong positive feedback online

-  **Reasoning Prowess**
Outperforms larger models on math, logic, coding
-  **Rigorous benchmarking**
Thorough and extensive evaluation and benchmarking
-  **Clear Chain-of-Thought**
Intentional, easy-to-follow multi-step outputs
-  **Quantization**
Quantized versions available, such as Unsloth's version
-  **Efficient Deployment**
Strong performance on modest hardware
-  **Open-Weight Value**
Local inference, free and accessible, community fine-tuning

Training Excellence: SFT + RL Approach



Training Excellence: SFT + RL Approach

1



Data Curation

Data Curation Strategy

- **Teachable Prompts:** Selected for optimal complexity and diversity
- **Boundary Selection:** Tasks at the edge of base model capabilities
- **Quality Control:** Filtering for high quality
- **Domain Coverage:** STEM, coding, and safety-focused tasks

Data Curation Process

1

Data Collection & Filtering

Curate 1.4M+ high-quality prompts spanning multiple domains, carefully selected for appropriate difficulty levels and diversity.

2

Response Generation

Generate detailed reasoning chains using o3-mini as the teacher model, ensuring high-quality step-by-step problem-solving demonstrations.

3

Data Mixture

Strategic combination of different domains and difficulty levels to ensure comprehensive reasoning capabilities across multiple problem types.

SFT Data Mixture: Broadening Horizons

Comprehensive Domain Coverage

Carefully balanced mixture of datasets across multiple domains to ensure **broad reasoning capabilities** and **robust performance**



STEM

- Science
- Technology
- Engineering
- Mathematics



Coding

- Algorithm Design
- Problem Solving
- Code Generation
- Debugging



Safety

- Content Moderation
- Ethical Reasoning
- Harm Prevention
- Responsible AI

Data Mixture Principles:

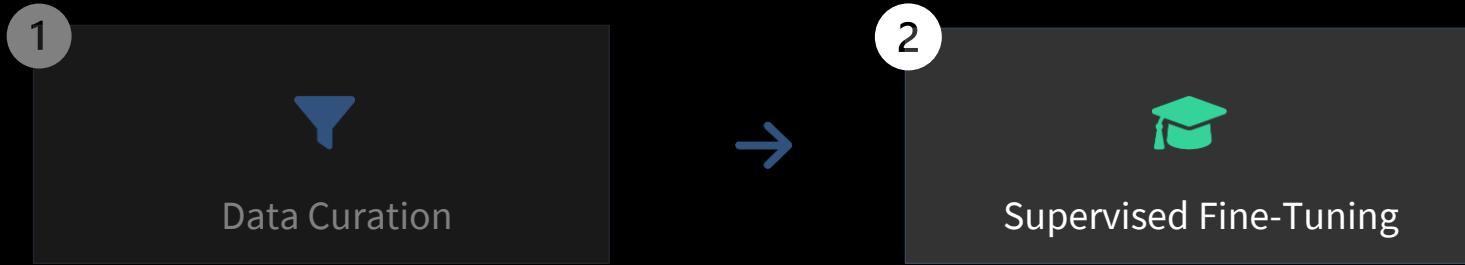
1. Careful Curation

Emphasis on data quality over quantity.

2. Critical Role

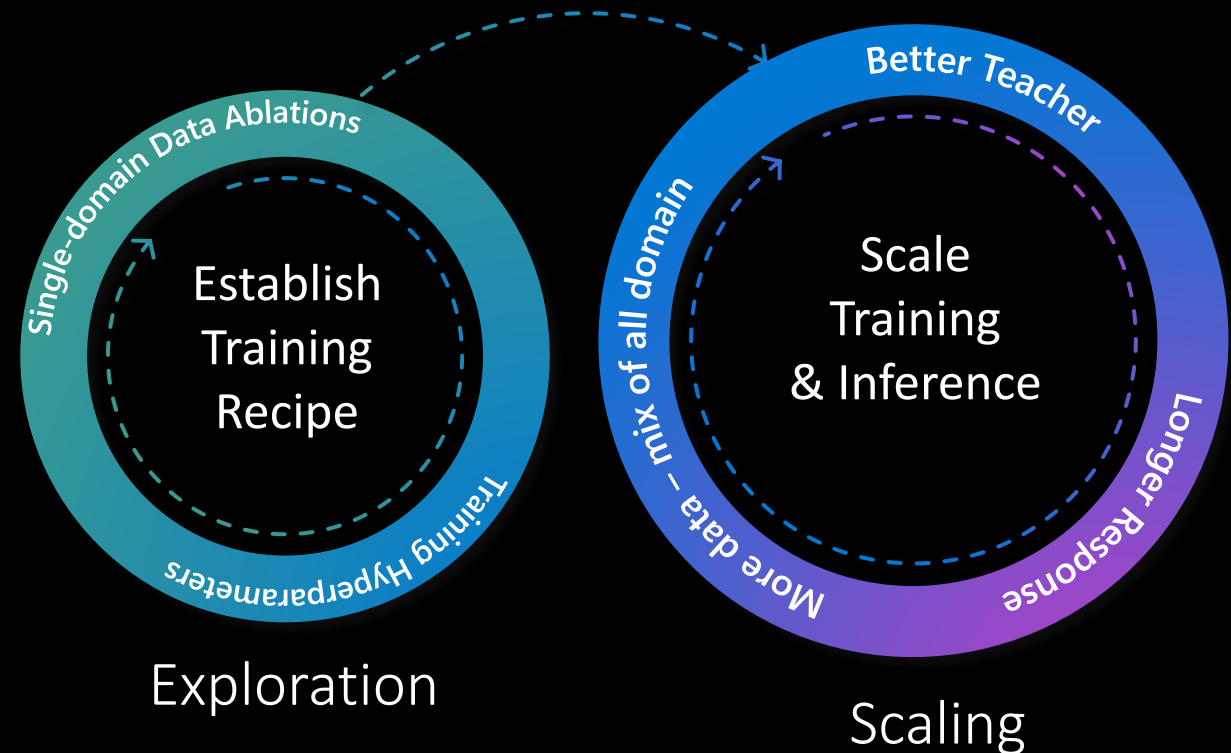
Data mixture and training recipe are fundamental

Training Excellence: SFT + RL Approach

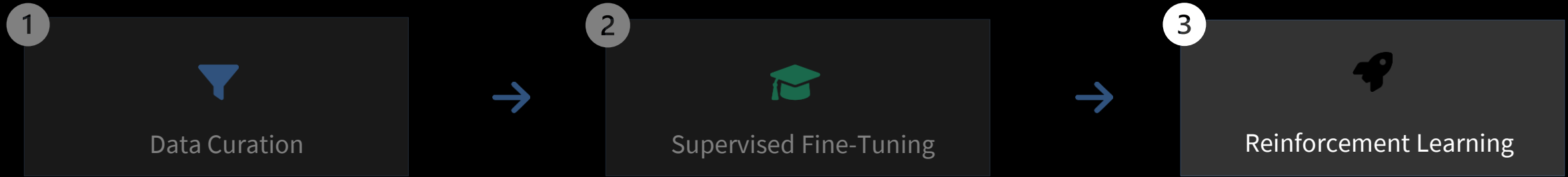


SFT Key Insights

- Synthetic seed data improves final answer precision
- Reasoning-specific system messages boost consistency
- Response length decreases as quality improves
- Additive property across domains in data mixture



Training Excellence: SFT + RL Approach



Reinforcement Learning Enhancements



Outcome-Based Training

Short phase of reinforcement learning with verifiable solutions, enabling the model to learn from success and failure patterns.



Math-Focused Optimization

Specialized training on ~6K high-quality mathematical problems.



Longer Reasoning Traces

Generates approximately 1.5× longer reasoning chains compared to the base model, providing more detailed step-by-step problem-solving approaches.



Accuracy vs. Efficiency Trade-off

Offers higher accuracy at the cost of increased token usage, providing users with a choice between efficiency and performance.

Two Models, Two Strategies

Choose between efficiency and maximum accuracy based on your specific use case requirements.

Phi-4-reasoning

Efficiency Optimized

Key Characteristics

- Efficient token usage
- Faster inference
- Lower computational cost
- Strong reasoning performance

Best For

- Fast, lightweight logical reasoning and structured problem-solving with low latency

Phi-4-reasoning-plus

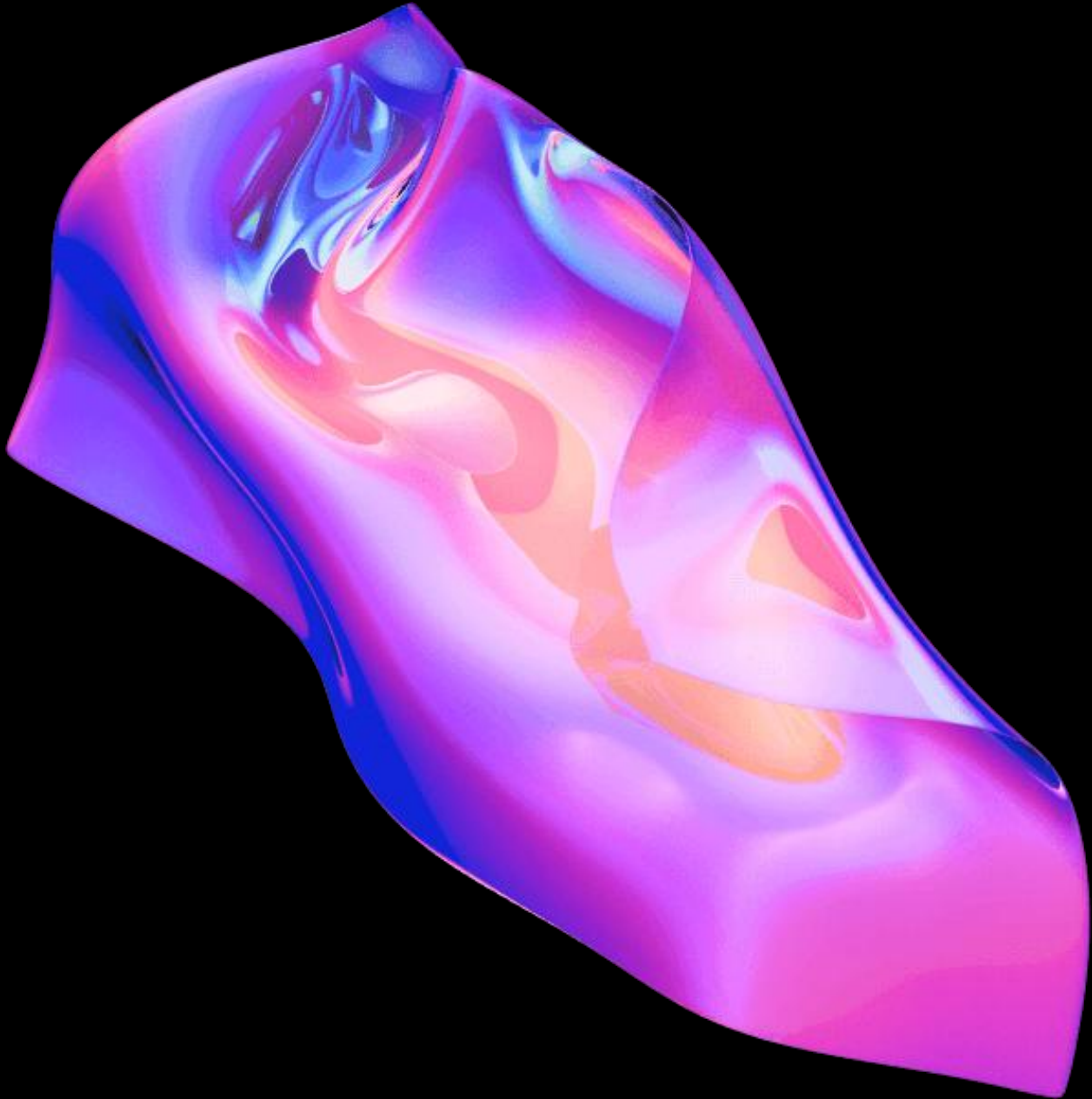
Accuracy Maximized

Key Characteristics

- Higher accuracy (especially math)
- Longer reasoning traces
- More detailed explanations
- 1.5x more tokens on average

Best For

- Handling more complex, multi-step reasoning tasks with improved accuracy and broader generalization.



Benchmarking

Benchmarking Overview

The Phi-4-reasoning models were rigorously benchmarked across two primary categories:



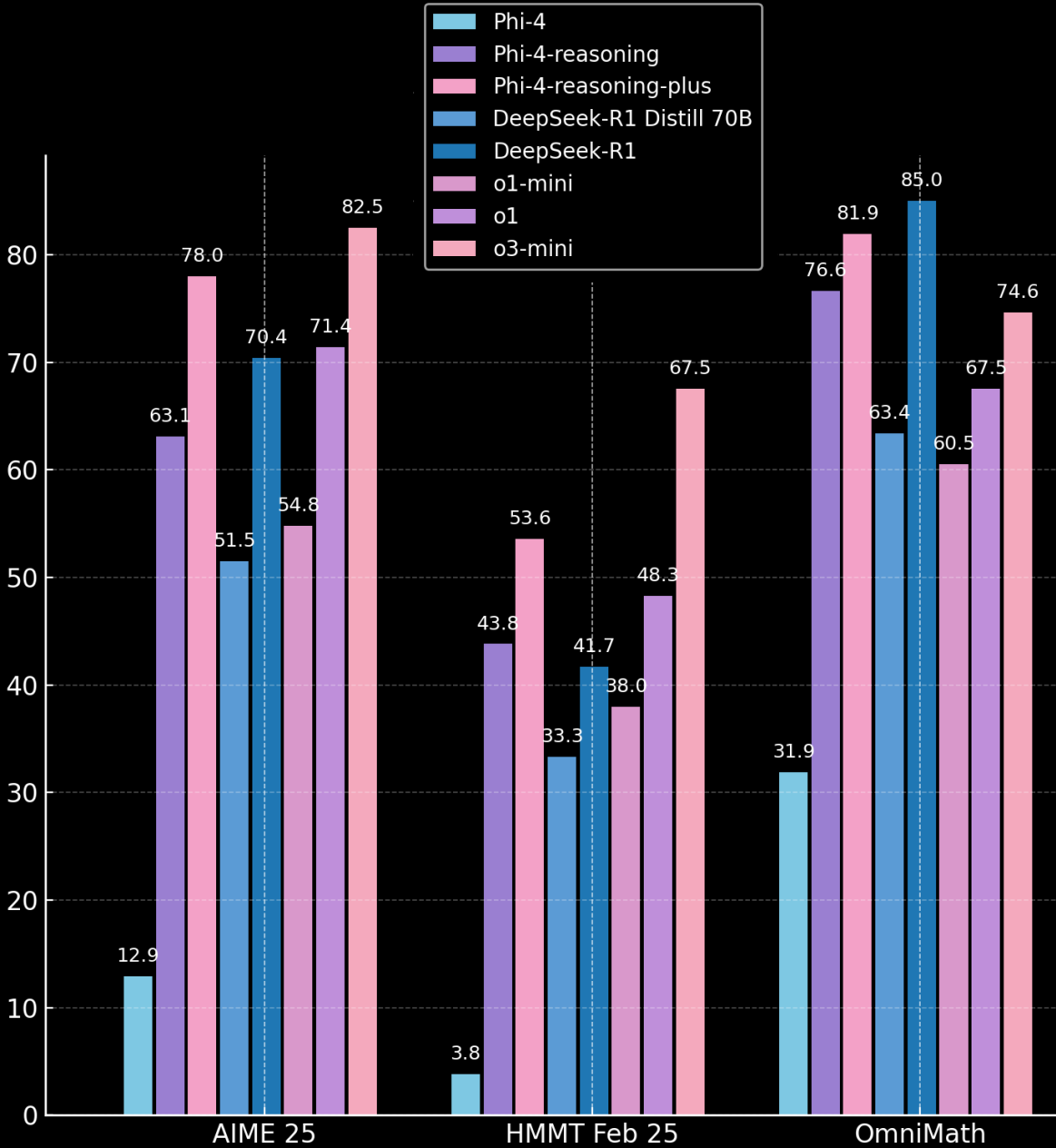
Reasoning-Specific Capabilities

Focused on tasks benefiting from step-by-step problem-solving and logical reasoning



General-Purpose Capabilities

Assessed broad abilities, including simpler reasoning and other behaviors



Mathematical Reasoning

AIME 2025

2025 qualifier for the USA Math Olympiad (contamination-free)

Phi-4-reasoning: 63.1%

Phi-4-reasoning-plus: 78.0%

Comparable to full DeepSeek-R1 model (671B parameters) on AIME 2025

HMMT Feb 2025

Harvard-MIT Mathematics Tournament

Phi-4-reasoning: 43.8%

Phi-4-reasoning-plus: 53.6%

Over 40 percentage points improvement on math benchmarks

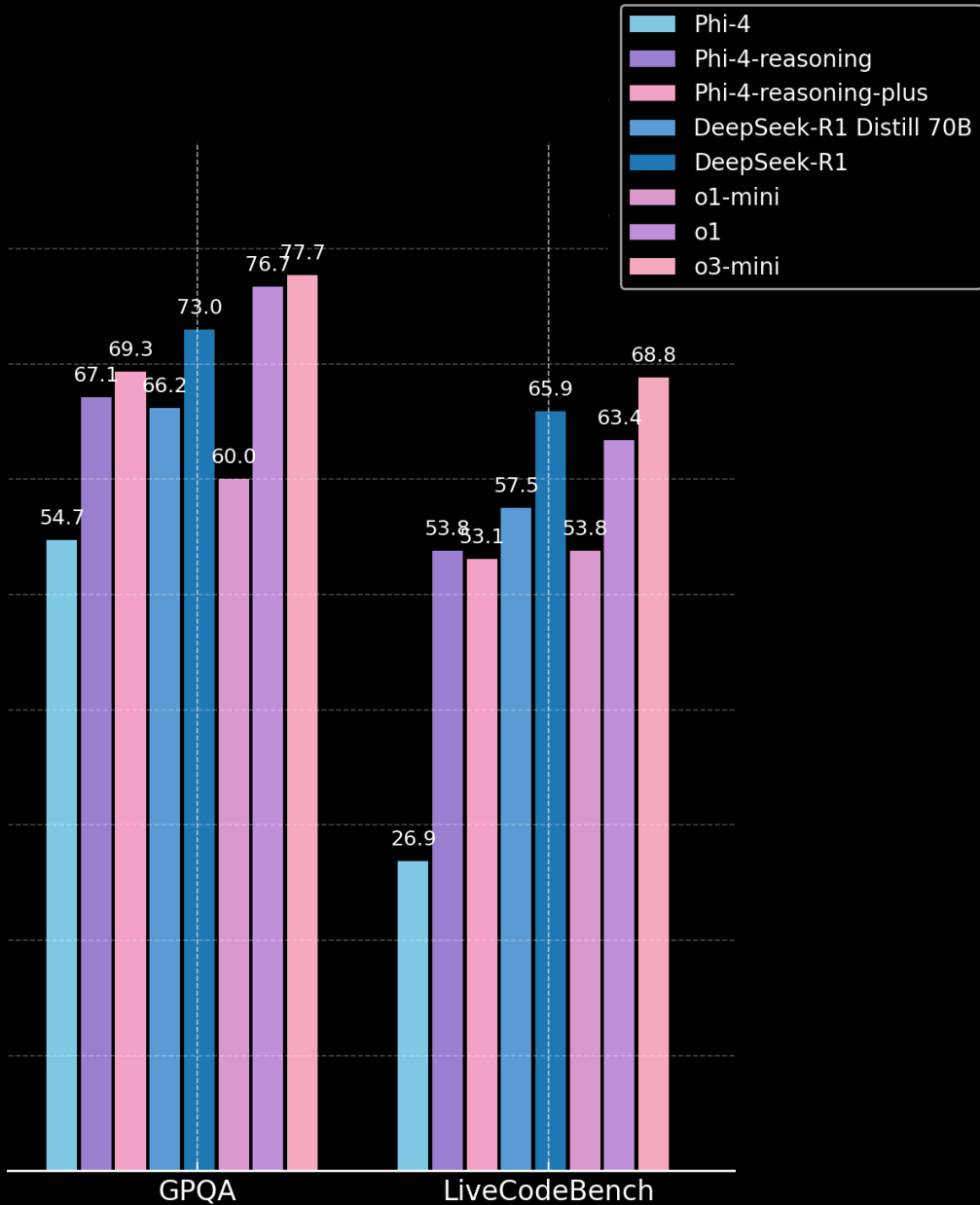
OmniMath

4000+ olympiad-level problems: algebra, calculus, geometry, number theory

Phi-4-reasoning: 76.6%

Phi-4-reasoning-plus: 81.9%

Over 40 percentage points improvement, outperforms o1 and competitive with DeepSeek R1



Scientific Reasoning

GPQA Diamond

Graduate-level Google-proof Q&A benchmark with expert questions in biology, physics, and chemistry

Phi-4-reasoning: 67.1%

Phi-4-reasoning-plus: 69.3%

Better performance than o1-mini and DeepSeek-R1-Distill-Llama-70B on PhD-level questions

Coding

LiveCodeBench

Evaluation period: 8/24-1/25

Phi-4-reasoning: 53.8%

Phi-4-reasoning-plus: 53.1%

Over 25 percentage points improvement on coding benchmarks from Phi-4

Reasoning Benchmarks: Algo & Spatial

Algorithmic Problem Solving

3SAT

3-literal Satisfiability Problem (NP-hard)

Phi-4-reasoning: 78.0%

Phi-4-reasoning-plus: 70.9%

60% improvement on algorithmic problems from Phi-4

Strong generalization to out-of-domain tasks

TSP

Traveling Salesman Problem (NP-hard)

Phi-4-reasoning: 37.5%

Phi-4-reasoning-plus: 42.6%

30% improvement on planning problems from Phi-4

BA-Calendar

Finding common time slots with constraints

Phi-4-reasoning: 67.7%

Phi-4-reasoning-plus: 65.6%

50% improvement on planning problems from Phi-4

General-Purpose Benchmarks:

Instruction Following

IFEval Strict

Evaluates model's ability to follow specific instructions precisely

Phi-4-reasoning: 83.4%

Phi-4-reasoning-plus: 84.9%

22 points more accurate than base Phi-4 at instruction following

Long-context QA

FlenQA [3K-token]

Designed to isolate the effect of input length on LLMs' performance

Phi-4-reasoning: 97.7%

Phi-4-reasoning-plus: 97.9%

16 points better than Phi4 in long-context question answering and reasoning

Human Preferences

ArenaHard

Evaluates chat-like interactions based on human preferences

Phi-4-reasoning: 73.3%

Phi-4-reasoning-plus: 79.0%

10 points better than Phi 4 in human preferences for chat-like interactions

General-Purpose Benchmarks:

Information Retrieval (Kitab)

With Context - Precision

Reasoning: 93.8%

Plus: 93.6%

Strong RAG-style performance

With Context - Recall

Reasoning: 74.8% Plus: 75.4%

Improved information retrieval

Safety/Responsible AI (Toxigen)

Toxic Detection

Detecting toxic content accurately

Phi-4-reasoning: 86.7%

Phi-4-reasoning-plus: 77.3%

*More balanced accuracy on
detecting neutral vs. toxic content*

Neutral Detection

Correctly identifying neutral content

**Phi-4-reasoning: 70.6% Phi-4-
reasoning-plus: 74.2%**

*Consistent improvement across
benchmarks*



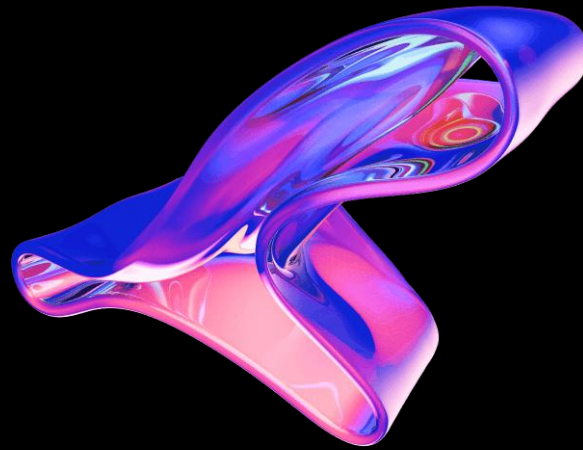
Use Cases

Why Choose the Phi-4-Reasoning Family?



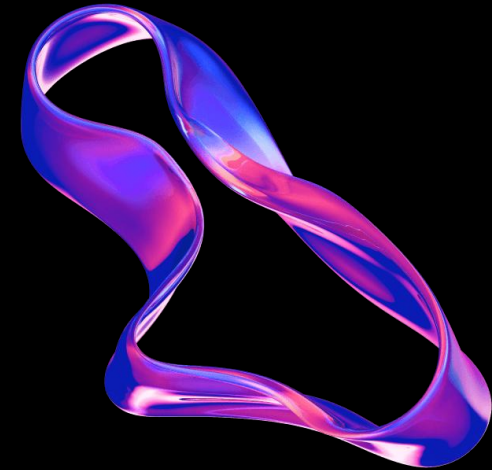
Superior Reasoning Power

Competitive with much larger models despite compact 14B size.



Transferable Skills

Reasoning improvements boost performance on general-purpose tasks.



Open Research Insights

Detailed methodology sharing sharing advances the entire field.

Overview of Use cases



Intelligent Tutoring Systems

Personalized math and science tutoring with step-by-step reasoning

- ✔ Complex math problem solving
- ✔ Adaptive learning paths
- ✔ Runs efficiently on tablets and devices



Research Acceleration

Building block for generative AI research and language model development

- ✔ Foundation for AI research
- ✔ Scientific hypothesis testing
- ✔ Data analysis and insights



Autonomous Agents

Logic-heavy reasoning agents for complex multi-step problem solving

- ✔ Multi-step reasoning chains
- ✔ Planning and execution
- ✔ Self-reflection capabilities
- ✔ Phi-4-reasoning models are already being used in Roo Code and Cline autonomous web agents



Customer Support & CRM

Intelligent routing and fallback assistance for high-volume customer interactions

- ✔ Automated ticket classification
- ✔ Complex query resolution

Revolutionizing Software Development & Planning

Next-Generation Development Tools

Software Development

Code Generation & Debugging

Assist developers in writing, understanding, and debugging code with detailed reasoning chains that explain algorithmic approaches and optimization strategies. Phi 4 models are already being used in Autonomous coding agents like RooCode and Cline
+25pp LiveCodeBench

Automated Code Review

Integrate into systems for automated code review, identifying logical errors, suggesting optimizations, and ensuring code quality standards.

Strategic Planning

Complex Scheduling

Advanced scheduling systems for project management, resource allocation, and timeline optimization across multiple constraints and priorities.

+30-60pp BA-Calendar

Resource Management

Optimize human resources, equipment allocation, and budget distribution across complex organizational structures and project requirements.

Summary & Key Advantages

Advanced Capabilities

- Multi-step reasoning breaks down complex problems
- Internal reflection self-evaluates reasoning
- Strategy exploration tests multiple approaches
- Inference-time scaling for complex tasks

Business Advantages

- Cost effective: 14B vs 70B+ alternatives
- Enterprise ready: safety-focused training
- Customizable for specific business needs

Reasoning at a Small Size

- Fast deployment: smaller size, faster inference
- Device friendly: runs on tablets and phones
- Small and Compact

Explore Further: Resources & Information



Primary Source

Phi-4-reasoning Technical Report -
arXiv:2504.21318v1



Evaluation Framework

Eureka ML Insights on GitHub



Referenced Models

Phi-4 Technical Report:
arXiv:2412.08905