# CHAIN AGENTS

## QUALITY

### SEAL

#### URL

<https://scale.com/leaderboard/tool_use>

#### Methodology

<https://scale.com/leaderboard/methodology>

#### Benchmarks

Score

## SPEED

## LATENCY

## COST

## CONTEXT WINDOW

# CHAT

## QUALITY

### Chatbot Arena

#### URL

<https://scale.com/leaderboard/tool_use>

#### Methodology

<https://scale.com/leaderboard/methodology>

#### Benchmarks

* Instruction Following
* Multi-Turn
* Longer Query
* Arena Score

#### *Artificial Analysis*

#### URL

<https://artificialanalysis.ai/models#quality>

#### Methodology

<https://artificialanalysis.ai/methodology>

#### Benchmarks

* Quality Index
* Reasoning & Knowledge (MMLU)
* Communication (LMSys Chatbot Arena ELO Score)

### Hugging Face

#### URL

<https://huggingface.co/spaces/open-llm-leaderboard/open_llm_leaderboard>

#### Methodology

<https://huggingface.co/docs/leaderboards/open_llm_leaderboard/about>

#### Benchmarks

* IFEval
* MUSR

### Vellum

#### URL

<https://www.vellum.ai/llm-leaderboard>

#### Methodology

<https://www.vellum.ai/llm-leaderboard>

#### Benchmarks

* Reasoning

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# SOLVE COMPLEX PROBLEMS

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# GENERATE CODE

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# GENERATE IMAGES

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# GENERATE TEXT

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# GENERATE VIDEO

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# SOLVE MATH PROBLEMS

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# CONVERT SPEECH TO TEXT

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

# CONVERT TEXT TO SPEECH

## QUALITY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## SPEED

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## LATENCY

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## COST

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

## CONTEXT WINDOW

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks

### [Leaderboard]

#### URL

#### Methodology

#### Benchmarks