OpenSeeker: First Fully Open-Source Search Agent Beats Industry Giants with 11,700 Training Samples

Kabui, Charles

2026-03-18

Researchers at Shanghai Jiao Tong University released OpenSeeker, the first search agent to open-source both model weights and 100% of its training data. Built on Qwen3-30B-A3B (a 30B-parameter mixture-of-experts model with 3B active parameters), it was fine-tuned

on just 11,700 synthesized samples using a single round of supervised fine-tuning, with no reinforcement learning or continual pre-training. Two techniques make the data work: fact-grounded QA synthesis, which reverse-engineers the web graph to generate complex multi-hop questions, and denoised trajectory synthesis, which uses retrospective summarization to clean up the messy HTML context that web-browsing agents encounter. The result: 48.4% on BrowseComp-ZH, beating Alibaba's Tongyi DeepResearch (46.7%) despite Alibaba using a far heavier pipeline of continual pre-training, SFT, and RL. On English benchmarks, it scores 29.5% on BrowseComp, 74.0% on xbench, and 59.4% on WideSearch, all state-of-the-art for open models of comparable size.

The real competitive moat in search agents has never been model weights. It is the high-quality browsing data that only companies like Perplexity, Google, and Alibaba can collect at scale. OpenSeeker proves that carefully designed synthetic data, 11,700 samples versus millions, can match that advantage. Any researcher can now download the full training dataset and build a competitive search agent without corporate-scale resources or proprietary web-browsing logs.

This follows a broader pattern: open-source models are closing the gap with proprietary ones not just in general chat, but in specialized agentic tasks. For previous coverage of open models reaching frontier performance, see GLM-5.

Sources:

- OpenSeeker Paper (arXiv)
- OpenSeeker GitHub Repository
- OpenSeeker-v1-30B-SFT Model (HuggingFace)
- OpenSeeker-v1 Training Data (HuggingFace)
- HuggingFace Daily Papers, March 17 (#3 Paper of the Day)