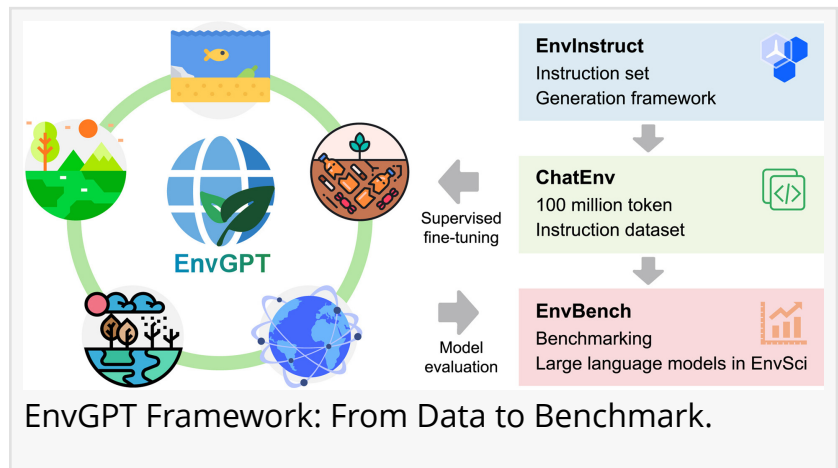


EnvGPT: a specialized AI tool for climate, water, and soil science challenges

GA, UNITED STATES, September 12, 2025 /EINPresswire.com/ -- Large language models (LLMs) are transforming specialized fields, yet environmental science—with its complex terminology and interdisciplinary nature—has lagged behind. This study introduces a unified framework to fine-tune an 8-billion-parameter model, [EnvGPT](#), using a carefully curated instruction dataset spanning climate change, ecosystems, water resources, soil management, and renewable energy. The model achieves state-of-the-art performance, rivaling much larger models in accuracy and relevance, offering a scalable solution for environmental research and policy support.



Environmental science integrates diverse disciplines like ecology, hydrology, and climate science, requiring models that understand specialized jargon and heterogeneous data. While general-purpose Large language models (LLMs) have advanced fields like medicine and law, they struggle with domain-specific environmental tasks due to limited training on relevant corpora. Previous efforts like ClimateGPT and WaterGPT focused on narrow subdomains, lacking a unified, cross-disciplinary approach. Based on these challenges, there is a critical need to develop integrated frameworks that generate high-quality environmental data and enable rigorous model evaluation.

Published on August 1, 2025, in *Environmental Science and Ecotechnology*, researchers from Southern University of Science and Technology and Tsinghua University unveiled EnvGPT—a fine-tuned language model specifically designed for environmental science. The study presents a comprehensive pipeline including a multi-agent instruction generator (EnvInstruct), a balanced 100-million-token dataset (ChatEnv), and a 4998-item benchmark (EnvBench) to train and evaluate the model across five core environmental themes.

The research team constructed EnvCorpus from open-access environmental journals, covering five key themes, and used a multi-agent GPT-4 system to generate 112,946 instruction-response

pairs. EnvGPT was fine-tuned using low-rank adaptation (LoRA), significantly reducing computational cost while maintaining performance. On the independently designed EnvBench, EnvGPT outperformed similarly sized models like LLAMA-3.1-8B and Vicuna-1.5-7B, and even matched the performance of the much larger Qwen2.5-72B and closed-source GPT-4o-mini in factual accuracy and relevance. Notably, it achieved 92.06% accuracy on the EnviroExam benchmark—a test based on university-level multiple-choice questions—surpassing baseline models by ~8 points. The model also excelled in real-world applicability, especially in interdisciplinary and complex reasoning tasks, as validated by the ELLE dataset.

"This work demonstrates how targeted fine-tuning with domain-specific data can elevate compact models to compete with giants in the field. EnvGPT sets a new standard for AI applications in environmental science," said Dr. Qing Hu, corresponding author and lead researcher at the State Key Laboratory of Soil Pollution Control and Safety.

EnvGPT can support researchers, educators, and policymakers by providing accurate, domain-aware responses to complex environmental queries. The open release of ChatEnv and EnvBench enables reproducible research and encourages community-driven improvements. Future work may integrate retrieval-augmented generation and multimodal data to enhance real-time reasoning and keep pace with evolving scientific knowledge.

References

DOI

[10.1016/j.jese.2025.100608](https://doi.org/10.1016/j.jese.2025.100608)

Original Source URL

<https://doi.org/10.1016/j.jese.2025.100608>

Funding information

This research was supported by the National Key Research and Development Program of China (2024YFC3711800) and the High-level University Special Fund (G03050K001).

Lucy Wang

BioDesign Research

[email us here](#)

This press release can be viewed online at: <https://www.einpresswire.com/article/848497991>

EIN Presswire's priority is source transparency. We do not allow opaque clients, and our editors try to be careful about weeding out false and misleading content. As a user, if you see something we have missed, please do bring it to our attention. Your help is welcome. EIN Presswire, Everyone's Internet News Presswire™, tries to define some of the boundaries that are reasonable in today's world. Please see our Editorial Guidelines for more information.

© 1995-2025 Newsmatics Inc. All Right Reserved.