

Researchers are developing AI to make the internet more accessible

January 9 2024, by Tatyana Woodall

(a) Find one-way flights from New York to Toronto.

(b) Book a roundtrip on July 1 from Mumbai to London and vice versa on July 5 for two adults.

(c) Find a flight from Chicago to London on 20 April and return on 23 April.

(d) Find Elon Musk's profile and follow, start notifications and like the latest tweet.

(e) Browse comedy films streaming on Netflix that was released from 1992 to 2007.

(f) Open page to schedule an appointment for car knowledge test.

Domain	Percentage
Travel	27.4%
Info	20.4%
Service	18.3%
Shopping	17.5%
Entertainment	16.1%
Other	6.2%
Airlines	5.8%
Health	6.5%
Speciality	5.0%
Restaurant	4.3%
Government	3.4%
Home Service	2.5%
Ground	3.7%
Car Rental	2.4%
Hotel	1.8%
Finance	2.7%
Cooking	2.9%
Weather	1.9%
Pet	2.1%
Shipping	1.8%
Moving	1.7%
Auto	3.4%
Digital	2.3%
General	3.3%
Department	1.6%
Fashion	1.6%
Event	3.8%
Music	3.5%
Sports	3.5%
Movie	3.2%
Game	2.0%

Sample tasks and all domains featured in MIND2WEB. The array of diversity allows for testing an agent's generalizability across tasks on the same website (a vs. b), similar tasks on different websites (a vs. c), and even to entirely disparate tasks, websites, and domains (d–f). Credit: *arXiv* (2023). DOI:

10.48550/arxiv.2306.06070

In an effort to make the internet more accessible for people with disabilities, researchers at The Ohio State University have begun developing an artificial intelligence agent that could complete complex tasks on any website using simple language commands.

In the three decades since it was first released into the [public domain](#), the world wide web has become an incredibly intricate, dynamic system. Yet because internet function is now so integral to society's well-being, its complexity also makes it considerably harder to navigate.

Today there are billions of websites available to help access information or communicate with others, and many tasks on the internet can take more than a dozen steps to complete. That's why Yu Su, co-author of the study and an assistant professor of computer science and engineering at Ohio State, said their work, which uses information taken from live sites to create web agents—online AI helpers—is a step toward making the [digital world](#) a less confusing place.

"For some people, especially those with disabilities, it's not easy for them to browse the internet," said Su. "We rely more and more on the computing world in our daily life and work, but there are increasingly a lot of barriers to that access, which, to some degree, widens the disparity."

The study was presented in December at the Thirty-seventh Conference on Neural Information Processing Systems (NeurIPS), a flagship conference for AI and machine learning research. It is [available](#) on the *arXiv* preprint server.

By taking advantage of the power of large language models, the agent works similarly to how humans behave when browsing the web, said Su. The Ohio State team showed that their model was able to understand the layout and functionality of different websites using only its ability to process and predict language.

Researchers started the process by creating Mind2Web, the first dataset for generalist web agents. Though previous efforts to build web agents focused on toy simulated websites, Mind2Web fully embraces the complex and dynamic nature of real-world websites and emphasizes an agent's ability of generalizing to entirely new websites it has never seen before.

Su said that much of their success is due to their agent's ability to handle the internet's ever-evolving learning curve. The team lifted over 2,000 open-ended tasks from 137 different real-world websites, which they then used to train the agent.

Some of the tasks included booking one-way and round-trip international flights, following celebrity accounts on Twitter, browsing comedy films from 1992 to 2017 streaming on Netflix, and even scheduling car knowledge tests at the DMV. Many of the tasks were very complex—for example, booking one of the [international flights](#) used in the model would take 14 actions. Such effortless versatility allows for diverse coverage on a number of websites, and opens up a new landscape for future models to explore and learn in an autonomous fashion, said Su.

"It's only become possible to do something like this because of the recent development of large language models like ChatGPT," said Su. Since the chatbot became public in November 2022, millions of users have used it to automatically generate content, from poetry and jokes to cooking advice and medical diagnoses.

Still, because one [website](#) could contain thousands of raw HTML elements, it would be too costly to feed so much information to a single large language model. To address this gap, the study also introduces a framework called MindAct, a two-pronged agent that uses both small and large language models to carry out these tasks. The team found that by using this strategy, MindAct significantly outperforms other common modeling strategies and is able to understand various concepts at a decent level.

With more fine-tuning, the study points out, the model could likely be used in tandem with both open-and closed-source large language models such as Flan-T5 or GPT-4. However, their work does highlight an increasingly relevant ethical problem in creating flexible artificial intelligence, said Su. While it could certainly serve as a helpful agent to humans surfing the web, the model could also be used to enhance systems like ChatGPT and turn the entire internet into an unprecedentedly powerful tool, said Su.

"On the one hand, we have great potential to improve our efficiency and to allow us to focus on the most creative part of our work," he said. "But on the other hand, there's tremendous potential for harm." For instance, autonomous agents able to translate online steps into the real world could influence society by taking potentially dangerous actions, such as misusing financial information or spreading misinformation.

"We should be extremely cautious about these factors and make a concerted effort to try to mitigate them," said Su. But as AI research continues to evolve, he notes that it's likely society will experience major growth in the commercial use and performance of generalist web agents in the years to come, especially as the technology has already gained so much popularity in the public eye.

"Throughout my career, my goal has always been trying to bridge the gap

between human users and the computing world," said Su. "That said, the real value of this tool is that it will really save people time and make the impossible possible."

More information: Xiang Deng et al, Mind2Web: Towards a Generalist Agent for the Web, *arXiv* (2023). [DOI: 10.48550/arxiv.2306.06070](https://doi.org/10.48550/arxiv.2306.06070)

Provided by The Ohio State University

Citation: Researchers are developing AI to make the internet more accessible (2024, January 9) retrieved 17 April 2024 from <https://techxplore.com/news/2024-01-ai-internet-accessible.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.