Researchers have a magic tool to understand AI: Harry Potter

MORE THAN two decades after J.K. Rowling introduced the world to a universe of magical creatures, forbidden forests, and a teenage wizard, Harry Potter is finding renewed relevance in a very different body of literature: AI research.

A growing number of researchers are using the best-selling Harry Potter books to experiment with generative artificial intelligence (AI), technology, citing the series' enduring influence in popular culture and the wide range of language data and complex wordplay within its pages. Reviewing a list of studies and academic papers referencing Harry Potter offers a snapshot into cutting-edge AI research — and some of the thorniest questions facing the technology.

In perhaps the most notable recent example, Harry, Hermione, and Ron star in a paper titled "Who's Harry Potter?" that sheds light on a new technique helping large language models to selectively forget information. It's a high-stakes task for the industry: Large language models, which power AI chatbots, are built on vast amounts of online data, including copyrighted material and other problematic content. That has led to lawsuits and public scrutiny for some AI companies.

The paper's authors, Microsoft researchers Mark Russinovich and Ronen Eldan, said they've demonstrated that AI models can be altered or edited to remove any knowledge of the existence of the Harry Potter books, including characters and plots, without sacrificing the AI system's overall decision-making and analytical abilities.

The duo said they chose the books because of their universal familiarity. "We believed that it would be easier for people in the research community to evaluate the model resulting from our technique and confirm for themselves that the content has indeed been 'unlearned,'" said Mr. Russinovich, chief technology officer of Microsoft Azure. "Almost anyone can come up with prompts for the model that would probe whether or not it 'knows' the books. Even people who haven't read the books would be aware of plot elements and characters."

In another study, researchers from the University of Washington in Seattle, University of California at Berkeley, and the Allen Institute for AI developed a new language model called Silo that can remove data to reduce legal risks. However, the model's performance significantly dropped if trained only on low-risk text such as out-of-copyright books or government documents, they said in a paper published earlier this year.

To go deeper, the researchers used Harry Potter books to see if individual pieces of text influence an AI system's performance. They created two datastores, or collections of websites and documents. The first included all published books except the first Harry Potter book; another included all books in the series but the second, and so on. "When the Harry Potter books are removed from the datastore, the perplexity gets worse," the researchers said, referring to the measure of accuracy of AI models.

— *Bloomberg*

FULL STORY

Read the full story by scanning the QR code or by typing the link <tinyurl.com/yx3sz4kf>

Shared data centers seen to help individuals, companies compute 'true' cost of power

DATA CENTERS shared by energy firms can leverage high-level computing to break down the price of electricity and improve consumer awareness of its true cost, an industry player said.

Energies PH, Inc. (EPHI) is looking to launch the Energy Community Digital Platform next year, which consists of Tier 3 colocation data centers for energy firms and use blockchain algorithms to accurately measure electricity consumption, cost, and efficiency, said Antonio A. Ver, EPHI chief executive officer.

"It's a community of players in the industry that is striving to go digital and compute the real price of electricity," Mr. Ver said about EPHI in an interview with *BusinessWorld*.

Their digital platform will help improve power price discovery, he said, as it will also allow people to verify their own meter readings via a website.

"For the benefit of everyone in the marketplace, nobody will be overpriced and you can know the correct amount of your electricity use," he said.

"Consumers will be given access to the data produced by the data center," Mr. Ver added.

He said he expects EPHI's initiative to empower the ecosystem against the rising prices of electricity in the country, fostering awareness about how much one should actually pay for energy.

"Sellers and buyers will be very careful," he said. "People will also be more comfortable buying renewable energy because they will pay right amount."

Firms and small electric cooperative resellers can also have a reliable source of prices through shared data centers, he added.

ALEXANDER JAWFOX/UNSPLASH

EPHI's platform includes initial sites in Baras, Rizal and Jaro, Iloilo, with plans to build three more in the second half of next year, including a data center complex in the Mambulao Special Economic Zone, Camarines Norte to cater South Luzon, Mr. Ver said.

He added that the company has already signed an agreement with a telco provider for the data centers' broadband and fiber networks powered by dense wavelength-division multiplexing.

Peering agreements are in place with tech giants such as Amazon, Google, and Microsoft, which Mr. Ver said will allow them to strategically expand operations in the country.

He cited as an example of a scheme akin to EPHI's Energy Community Digital Platform a partnership between Emirates Water and Electric Co. and Energy Exemplar, an optimization-based market simulation software provider for the energy sector, which was announced last week.

The collaboration aims to improve energy modeling capacity and integrated system planning through techno-economic cloud solutions, Mr. Ver noted.

— Miguel Hanz L. Antivola