

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

From hedge funds, to angel investors, to bull markets.

In the world of finance, names can sometimes be a bit misleading.

Take this one for instance, LSEG, aka London Stock Exchange Group.

LSEG is in London, and they have a stock exchange, but that's just part of what they do.

Today they connect the news, information, insights, and systems that make the global financial markets work.

You see, others do some of what they do, but not everything they do.

LSEG.

Hi, I'm John Gertner. I'm a contributor to the New York Times Magazine, and I write about science and technology.

This week's Sunday Read is a story I wrote for the magazine about Wikipedia.

It's a story that explains how the 22-year-old wonky online encyclopedia we've all consulted at one point

is so central to building artificial intelligence right now.

So, over the last few years, computer scientists have been creating what are known as large language models,

which are the AI brains that power the chatbots like ChatGPT.

And in order to build a large language model, they needed to gather vast knowledge banks of information.

And I mean, it's sort of dizzying how much information we're talking about here.

Some models ingest upwards of a trillion words.

And it all comes from public sources like Wikipedia, or Reddit, or Google's patent database.

What makes Wikipedia special is not just that it's free and accessible, but also that it's very highly formatted.

It contains just a tremendous amount of factual information that's maintained by a community of about 40,000 active editors in the English language version alone.

The problem with these new AI chatbots is that their fundamental goal is to converse with the user with a kind of human fluency of language, but they're not built to regurgitate data or to really be precise.

So, whether you're trying to understand historical topics or political upheavals or pandemics, these bots greatly simplify the world in a way that's maybe not conducive at all to our best interests as human beings.

AI chatbots have even been known to hallucinate and conjure falsehoods from whole cloth.

And another problem is that if they're fed only on their own synthetic data, these systems essentially break down.

So, if we were to go to AI instead of Wikipedia to find information, to solve our problems, to answer questions,

what would happen in the future where our knowledge is factually unreliable?

As I reported this story, I read a lot of what are called community notes,

which are the logs of Wikipedia editor meetings that they transcribe and make public.

And in one recent meeting, editors shared their worries about AI.

What's it going to do to Wikipedia?

I remember reading the notes for this meeting and one line from an editor popped out at me.

We want a future where knowledge is created by humans.

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

And I thought, well, that's really the essence of it, isn't it?

Can we really choose at this point the future we want?

So, here's my article, Wikipedia's Moment of Truth, read by Brian Nishi.

In early 2021, a Wikipedia editor peered into the future and saw what looked like a funnel cloud on the horizon.

The rise of GPT-3, a precursor to the new chatbots from OpenAI.

When this editor, a prolific Wikipedia who goes by the handle Barkeep49 on the site, gave the new technology a try,

he could see that it was untrustworthy.

The bot would readily mix fictional elements, a false name, a false academic citation, into otherwise factual and coherent answers.

But he had no doubts about its potential.

I think AI's day of writing a high-quality encyclopedia is coming sooner rather than later, he wrote in Death of Wikipedia,

an essay that he posted under his handle on Wikipedia itself.

He speculated that a computerized model could, in time, displace his beloved website and its human editors,

just as Wikipedia had supplanted the encyclopedia Britannica, which, in 2012, announced it was discontinuing its print publication.

Recently, when I asked this editor, he asked me to withhold his name because Wikipedia editors can be the targets of abuse.

If he's still worried about his encyclopedia's fate, he told me that the newer versions made him more convinced that chat GPT was a threat.

It wouldn't surprise me if things are fine for the next three years, he said of Wikipedia.

And then, all of a sudden, in year 405, things drop off a cliff.

Wikipedia marked its 22nd anniversary in January.

It remains, in many ways, a throwback to the Internet's utopian early days,

when experiments with open collaboration anyone can write and edit for Wikipedia

had yet to see the digital terrain to multi-billion-dollar corporations and data miners

advertising schemers and social media propagandists.

The goal of Wikipedia, as its co-founder Jimmy Wales described it in 2004,

was to create a world in which every single person on the planet is given free access to the sum of all human knowledge.

The following year, Wales also stated, we help the Internet not suck.

Wikipedia now has versions in 334 languages and a total of more than 61 million articles.

It consistently ranks among the world's 10 most visited websites,

yet is alone among that select group whose usual leaders are Google, YouTube, and Facebook in eschewing the profit motive.

Wikipedia does not run ads except when it seeks donations,

and its contributors, who make about 345 edits per minute on the site, are not paid.

In seeming to repudiate capitalism's imperatives, its success can seem surprising, even mystifying.

Some Wikipedians remark that their endeavor works in practice, but not in theory.

Wikipedia is no longer an encyclopedia, or at least not only an encyclopedia.

Over the past decade, it has become a kind of factual netting that holds the whole digital world

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

together.

The answers we get from searches on Google and Bing, or from Siri and Alexa, how old is Joe Biden, or what is an ocean submersible, derive in part from Wikipedia's data having been ingested into their knowledge banks. YouTube has also drawn on Wikipedia to counter misinformation.

The new AI chatbots have typically swallowed Wikipedia's corpus too, embedded deep within their responses to queries is Wikipedia data and Wikipedia text, knowledge that has been compiled over years of painstaking work by human contributors. While estimates of its influence can vary, Wikipedia is probably the most important single source in the training of AI models.

Without Wikipedia, generative AI wouldn't exist, says Nicholas Vincent, who will be joining the faculty of Simon Fraser University in British Columbia this month, and who has studied how Wikipedia helps support Google searches and other information businesses.

Yet, as bots like chatGPT become increasingly popular and sophisticated, Vincent and some of his colleagues wonder what will happen if Wikipedia, outflanked by AI that has cannibalized it, suffers from disuse and dereliction.

In such a future, a death of Wikipedia outcome is perhaps not so far-fetched. A computer intelligence, it might not need to be as good as Wikipedia, merely good enough, is plugged into the web and seizes the opportunity to summarize source materials and news articles instantly, the way humans now do with argument and deliberation.

On a conference call in March that focused on AI's threats to Wikipedia, as well as the potential benefits, the editors' hopes contended with anxiety.

While some participants seemed confident that generative AI tools would soon help expand Wikipedia's articles and global reach, others worried about whether users would increasingly choose chatGPT, fast, fluent, seemingly oracular, over a wonky entry from Wikipedia.

A main concern among the editors was how Wikipedians could defend themselves from such a threatening technological interloper, and some worried about whether the digital realm had reached a point where their own organization, especially in its striving for accuracy and truthfulness, was being threatened by a type of intelligence that was both factually unreliable and hard to contain. One conclusion from the conference call was clear enough.

We want a world in which knowledge is created by humans. But is it already too late for that?

Back in 2017, the Wikimedia Foundation and its community of volunteers began exploring how the Encyclopedia and its sister sites, like Wikidata and Wikimedia Commons, with their offerings of free information and images, could evolve by the year 2030.

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

The plan was to ensure that the Foundation, the non-profit that oversees Wikipedia, could protect and share the world's information in perpetuity. One outcome of that 2017 effort, which included a year's worth of meetings, was a prediction that Wikimedia would become the essential infrastructure of the ecosystem of free knowledge. Another conclusion was that trends like online misinformation would soon require far more vigilance, and a research paper commissioned by the Foundation found that artificial intelligence was improving at a rate that could change the way that knowledge is gathered, assembled, and synthesized. For that reason, the rollout of chat GPT did not elicit surprise inside the Wikipedia community, though several editors told me they were shocked by the speed of its adoption, which needed just two months after its release in late 2022 to gain an estimated 100 million users. Despite its stodgy appearance, Wikipedia is more tech-savvy than casual users might assume. With a small group of volunteers to oversee millions of articles, it has long been necessary for highly experienced editors, often known as administrators, to use semi-automated software to identify misspellings and catch certain forms of intentional misinformation. And because of its open-source ethos, the organization has at times incorporated technology made freely available by tech companies or academics, rather than go through a lengthy and expensive development process on its own. We've had artificial intelligence tools and bots since 2002, and we've had a team dedicated to machine learning since 2017. Selena Deckelman, Wikipedia's chief technology officer, told me. They're extremely valuable for a semi-automated content review, and especially for translations. How Wikipedia uses bots and how bots use Wikipedia are extremely different, however. For years, it has been clear that fledgling AI systems were being trained on the site's articles as part of the process whereby engineers scrape the web to create enormous data sets for that purpose. In the early days of these models, about a decade ago, Wikipedia represented a large percentage of the scraped data used to train machines. The encyclopedia was crucial not only because it's free and accessible, but also because it contains a motherload of facts, and so much of its material is consistently formatted. In more recent years, as so-called large-language models, or LLMs, increased in size and functionality, these are the models that power chatbots like ChatGPT and Google Spard. They began to take in far larger amounts of information. In some cases, their meals added up to well over a trillion words. The sources included not just Wikipedia, but also Google's patent database, government documents, Reddit's Q&A corpus, books from online libraries, and vast numbers of news articles on the web. But while Wikipedia's contribution in terms of overall volume is shrinking, and even as tech companies have stopped disclosing what data sets go into their AI models,

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

it remains one of the largest single sources for LLMs.

Jesse Dodge, a computer scientist at the Allen Institute for AI in Seattle, told me that Wikipedia might now make up between 3 and 5% of the scraped data an LLM uses for its training.

Wikipedia, going forward, will forever be super valuable, Dodge points out, because it's one of the largest well-curated data sets out there.

There is generally a link, he adds, between the quality of data a model trains on and the accuracy and coherence of its responses.

In this light, Wikipedia might be seen as a sheep caught in the jaws of a wolfish technology marketplace.

A free site created in achingly good faith, sharing knowledges by nature and act of kindness, Wikipedia noted in 2017 on a page devoted to its strategic direction, is being devoured by companies whose objectives, like charging for subscriptions as OpenAI recently began doing for its latest model, don't jibe with its own.

Yet, the relationships are more complicated than they appear.

Wikipedia's fundamental goal is to spread knowledge as broadly and freely as possible by whatever means.

About 10 years ago, when site administrators focused on how Google was using Wikipedia, they were in a situation that presaged the advent of AI chatbots.

Google's search engine was able, at the top of its query results, to present Wikipedia's work to users all over the world, giving the Encyclopedia far greater reach than before, an apparent virtue.

In 2017, three academic computer scientists, Connor McMahon, Isaac Johnson and Brent Hecht, conducted an experiment that tested how random users would react if just part of the contributions made to Google's search results by Wikipedia were removed.

The academics perceived an extensive interdependence.

Wikipedia makes Google a significantly better search engine for many queries, and Wikipedia, in turn, gets most of its traffic from Google.

One upshot from the collision with Google and others who repurpose Wikipedia's content was the creation two years ago of Wikimedia Enterprise, a separate business unit that sells access to a series of application programming interfaces that provide accelerated updates to Wikipedia articles.

Depending on whom you ask, the Enterprise unit is either a more formalized way for tech companies to direct the equivalent of large charitable donations to Wikipedia,

Google now subscribes, and altogether the unit took in \$3.1 million in 2022, or a way for Wikipedia to recoup some of the financial value it creates for the digital world, and thus help fund its future operations.

Practically speaking, Wikipedia's openness allows any tech company to access Wikipedia at any time, but the APIs make new Wikipedia entries almost instantly readable.

This speeds up what was already a pretty fast connection.

Andrew Lee, a consultant who works with museums to put data about their collections on Wikipedia, told me he conducted an experiment in 2019 to see how long it would take for a new Wikipedia article

about a pioneering balloonist named Vera Simons to show up in Google search results.

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

He found the elapsed time was about 15 minutes.

Still, the close relationship between search engines and Wikipedia has raised some existential questions for the latter.

Ask Google, what is the Russian-Ukrainian war, and Wikipedia is credited with some of its material briefly summarized.

But what if that makes you less likely to visit Wikipedia's article, which runs to some 10,000 words and contains more than 400 footnotes?

From the point of view of some of Wikipedia's editors, reduced traffic will oversimplify our understanding of the world and make it difficult to recruit a new generation of contributors.

It may also translate into fewer donations.

In the 2017 paper, the researchers noted that visits to Wikipedia had indeed begun to decline, and the phenomenon they identified became known as the paradox of reuse.

The more Wikipedia's articles were disseminated through other outlets and media, the more imperiled was Wikipedia's own health.

With AI, this reuse problem threatens to become far more pervasive.

Aaron Hafaker, who led the machine learning research team at the Wikimedia Foundation for several years and who now works for Microsoft, told me that search engine summaries at least offer users links and citations and a way to click back to Wikipedia.

The responses from large language models can resemble an information smoothie that goes down easy but contains mysterious ingredients.

The ability to generate an answer has fundamentally shifted, he says, noting that in a chat GPT answer there is literally no citation and no grounding in the literature as to where that information came from.

He contrasts it with the Google or Bing search engines.

This is different. This is way more powerful than what we had before.

Almost certainly that makes AI both more difficult to contend with and potentially more harmful, at least from Wikipedia's perspective.

A computer scientist who works in the AI industry but is not permitted to speak publicly about his work told me that these technologies are highly self-destructive, threatening to obliterate the very content which they depend upon for training.

It's just that many people, including some in the tech industry, haven't yet realized the implications.

Can the living be killed by the dead?

This is the question Hercule Poirot must answer in *A Haunting in Venice*, his most terrifying case yet, but in a world of shadows and secrets who can be trusted.

Starring Kenneth Branagh, James Brunner, and James Brunner,

Kenneth Branagh, Jamie Dornan, Tina Fey, Kelly Riley, Michelle Yeoh, presented by 20th Century Studios, *A Haunting in Venice*, directed by Kenneth Branagh, rated PG-13,

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

experienced the movie event only in theaters September 15, tickets on sale now.

Hi, this is Eric Kim with New York Times Cooking.

As a recipe developer, I spend a lot of my time trying to come up with dishes that are quick, easy, but also very special.

For me, that means dishes like gochugaru salmon.

It's a crispy salmon filet with a salty, sweet glaze that bubbles up in candies.

I love cooking this because it only takes 20 minutes.

I developed this recipe when I was down in Georgia with my family.

It stars a Korean red pepper powder called gochugaru.

I love the way it blooms in the maple syrup and the rice vinegar.

If you don't have gochugaru, you should totally get some.

It's super versatile.

It's not just that it brings heat, but fruity sweetness as well.

You can get this recipe and so many more ideas on New York Times Cooking.

Visit [nytcooking.com](https://www.nytimes.com/nytcooking.com) to get inspired.

According to recent data from the Wikimedia Foundation,

about 80% of that cohort is male,

and about 75% of those from the United States are white,

which has led to some gender and racial gaps in Wikipedia's coverage.

And lingering doubts about reliability remain.

For a popular article that might have thousands of contributors,

Wikipedia is literally the most accurate form of information

ever created by humans.

Amy Brukman, a professor at the Georgia Institute of Technology, told me.

But Wikipedia's short articles can sometimes be hit or miss.

They could be total garbage, says Brukman,

who is the author of the recent book *Should You Believe Wikipedia?*

An erroneous fact on a rarely visited page may endure for months or years.

And there continues to exist the ever-present threat of vandalism

or tampering with an article.

In 2017, for instance, a photo of the Speaker of the House, Paul Ryan,

was added to the entry on invertebrates.

As a Wikipedia editor whose first name is Jade put it to me,

we have a number of, I would say, almost professional trolls

who must dedicate just about as much time to creating spam,

creating vandalism, harassing people,

as I dedicate to improving Wikipedia.

Several academics told me that whatever Wikipedia's shortcomings,

they view the encyclopedia as a consensus truth, as one of them put it.

It acts as a reality check in a society where facts are increasingly contested.

The truth is less about data points.

How old is Joe Biden?

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

Than about complex events like the COVID-19 pandemic, in which facts are constantly evolving, frequently distorted and furiously debated. The truthfulness quotient is raised by Wikipedia's transparency. Most Wikipedia entries include footnotes, links to source materials and lists of previous edits and editors, and experienced editors are willing to intercede when an article appears incomplete or lacks what Wikipedians call verifiability. Moreover, Wikipedia's guidelines insist that its editors maintain an NPOV, neutral point of view, or risk being overruled, or in the argo of wiki culture, reverted. And the site has a bent toward self-examination. You can find long disquisitions on Wikipedia that explore Wikipedia's own reliability. An entry on how Wikipedia has fallen victim to hoaxes, runs to more than 60 printed pages. As difficult as the pursuit of truth can be for Wikipedians, though, it seems significantly harder for AI chatbots. ChatGBT has become infamous for generating fictional data points or false citations known as hallucinations. Perhaps more insidious is the tendency of bots to oversimplify complex issues, like the origins of the Ukraine-Russia war, for example. One worry about generative AI at Wikipedia, whose articles on medical diagnoses and treatments are heavily visited, is related to health information. A summary of the March conference call captures the issue. We're putting people's lives in the hands of this technology. For example, people might ask this technology for medical advice. It may be wrong, and people will die. This apprehension extends not just to chatbots, but also to new search engines connected to AI technologies. In April, a team of Stanford University scientists evaluated four engines powered by AI, BingChat, Neva AI, Perplexity AI, and UChat, and found that only about half of the sentences generated by the search engines in response to a query could be fully supported by factual citations. We believe that these results are concerningly low for systems that may serve as a primary tool for information-seeking users, the researchers concluded, especially given their facade of trustworthiness. What makes the goal of accuracy so vexing for chatbots is that they operate probabilistically

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

when choosing the next word in a sentence.
They aren't trying to find the light of truth in a murky world.
These models are built to generate text that sounds
like what a person would say.
That's the key thing, Jesse Dodge says.
So they're definitely not dead.
So they're definitely not built to be truthful.
I asked Margaret Mitchell, a computer scientist
who studied the ethics of AI at Google
whether factuality should have been
a more fundamental priority for AI.
Mitchell, who says she was fired from the company
after criticizing the direction of its work,
Google says she was fired for violating the company's
security policies, said that most would find that logical.
This common-sense thing, shouldn't we work on making it factual
if we're putting it forward for fact-based applications?
Well, I think for most people who are not in tech,
it's like, why is this even a question?
But Mitchell said the priorities at the big companies,
now in frenzied competition with one another,
are concerned with introducing AI products
rather than reliability.
The road ahead will almost certainly lead to improvements.
Mitchell told me that she foresees AI companies
making gains in accuracy and reducing biased answers
by using better data.
The state of the art until now has just been
a laissez-faire data approach, she said.
You just throw everything in
and you're operating with a mindset where the more data you have,
the more accurate your system will be,
as opposed to the higher quality of data you have,
the more accurate your system will be.
Jesse Dodge, for his part,
points to an idea known as retrieval,
whereby a chatbot will essentially consult
a high-quality source on the web
to fact-check an answer in real time.
It would even cite precise links
as some AI-powered search engines now do.
Without that retrieval element, Dodge says,
I don't think there's a way to solve the hallucination problem.
Otherwise, he says, he doubts that a chatbot answer

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

can gain factual parity with Wikipedia or the Encyclopedia Britannica. Market competition might help prompt improvement, too. Owen Evans, a researcher at a nonprofit in Berkeley, California, who studies truthfulness in AI systems, pointed out to me that OpenAI now has several partnerships with businesses, and those firms will care greatly about responses achieving a high level of accuracy. Google, meanwhile, is developing AI systems to work closely with medical professionals on disease detection and diagnostics. There's just going to be a very high bar there, he adds, so I think there are incentives for the companies to really improve this. At least for now, AI companies are focusing on what they call fine-tuning when it comes to factuality. Sandini Argoal and Girish Sustray, researchers at OpenAI, the company that created ChatGPT, told me that their newer AI model, GPT-4, has made significant improvements over earlier models in what they called factual content. Those advances stem mainly from a process known as reinforcement learning with human feedback to help AI models differentiate between good and bad answers. But ChatGPT clearly has a way to go, both to fix hallucinations and to provide complex, multi-layered, and accurate answers to historical questions. When I asked Argoal whether OpenAI's systems could ever be completely accurate or offer 400 footnotes, she said that it was possible. But there might always exist a tension between a model's ambition to be factual and its efforts to be creative and fluent. As an AI developer, she explained, the goal was not for a Chat model to regurgitate data it had been trained on. Rather, it was to see patterns of knowledge it could relate to users in fresh conversational language. In the future, Sustray added, AI systems might interpret whether a query requires a rigorous factual answer or something more creative. In other words, if you wanted an analytical report with citations and detailed attributions,

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

the AI would know to deliver that.

And if you desired a sonnet about the indictment of Donald Trump, well, it could dash that off instead.

In late June, I began to experiment with a plug-in the Wikimedia Foundation had built for ChatGPT.

At the time, this software tool was being tested by several dozen Wikipedia editors and Foundation staff members, but it became available in mid-July on the OpenAI website for subscribers who want augmented answers to their ChatGPT queries.

The effect is similar to the retrieval process that Jesse Dodge surmises might be required to produce accurate answers. GPT-4's knowledge base is currently limited to data it ingested by the end of its training period in September 2021.

A Wikipedia plug-in helps the bot access information about events up to the present day.

At least in theory, the tool, lines of code that direct a search for Wikipedia articles that answer a ChatBot query, gives users an improved, combinatorial experience.

The fluency and linguistic capabilities of an AI ChatBot merged with the factuality and currency of Wikipedia.

One afternoon, Chris Albin, who's in charge of machine learning at the Wikimedia Foundation, took me through a quick training session.

Albin asked ChatGPT about the Titan submersible, operated by the company OceanGate, whose whereabouts during an attempt to visit the Titanic's wreckage were still unknown.

Normally, you get some response that's like, my information cut off is from 2021, Albin told me.

But in this case, ChatGPT, recognizing that it couldn't answer Albin's question, what happened with OceanGate's submersible, directed the plug-in to search Wikipedia and only Wikipedia for text relating to the question.

After the plug-in found the relevant Wikipedia articles, it sent them to the bot, which in turn read and summarized them, then spit out its answer.

As the responses came back, hindered by only a slight delay, it was clear that using the plug-in always forced ChatGPT to append a note, with links to Wikipedia entries, saying that its information was derived from Wikipedia, which was made by volunteers.

And this.

As a large language model, I may not have summarized Wikipedia accurately. But the summary about the submersible struck me as readable, well-supported and current,

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

a big improvement from a ChatGPT response that either mangled the facts or lacked real-time access to the Internet. Albin told me, it's a way for us to sort of experiment with the idea of what does it look like for Wikipedia to exist outside of the realm of the website, so you could actually engage in Wikipedia without actually being on wikipedia.com. Going forward, he said, his sense was that the plug-in would continue to be available, as it is now, to users who want to activate it, but that eventually there's a certain set of plug-ins that are just always on. In other words, his hope was that any ChatGPT query might automatically result in the ChatBots checking facts with Wikipedia and citing helpful articles. Such a process would probably block many hallucinations as well. For instance, because ChatBots can be deceived by how a question is worded, false premises sometimes elicit false answers. Or, as Albin put it, if you were to ask, during the first lunar landing, who were the five people who landed on the moon, the ChatBot wants to give you five names. Only two people landed on the moon in 1969, however. Wikipedia would help by offering the two names, Buzz Aldrin and Neil Armstrong, and in the event that ChatBot remained conflicted, it could say it didn't know the answer and link to the article. The plug-in still lets ChatGPT get creative, but in limited ways. The following week, when I asked it for updates about the Ocean Gate submersible, I got a three-paragraph rundown of how the tragedy unfolded, including the deaths of five passengers. Then I asked it to formulate its answer and five bullet points, which it did instantly. Could it then adapt those five bullet points, I asked, so that a seven- or eight-year-old could understand? Here's a simpler version, ChatGPT said instantly, and offered just what I asked for, noting that the Titan was a special underwater vehicle and its implosion was a sad event. It wasn't perfect. I told ChatGPT that its bullet points seemed to overlook how Stockton Rush, Ocean Gate's chief executive, had been criticized for ignoring safety standards. You raise a valid point, it responded. Here's a revised version that addresses your concern.

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

Its fix took only a few seconds.

Within the Wikipedia community, there is a cautious sense of hope that AI, if managed right, will help the organization improve rather than crash.

Selena Deckelman, the chief tech officer, expresses that perspective most optimistically.

What we've proven over 22 years now is, we have a volunteer model that is sustainable, she told me.

I would say there are some threats to it.

Is it an insurmountable threat?

I don't think so.

The longtime Wikipedia editor who wrote *Death of Wikipedia* told me that he feels there is a case to be made for a good outcome in the coming years, even if the longer term seems far less certain.

The Wikimedia plugin is the first significant move toward protecting its future.

Projects are also in the works to use recent advances in AI internally.

Albin says that he and his colleagues are in the process of adapting AI models that are off the shelf, essentially models that have been made available by researchers for anyone to freely customize so that Wikipedia's editors can use them for their work.

One focus is to have AI models aid new volunteers, say, with step-by-step chatbot instructions as they begin working on new articles, a process that involves many rules and protocols and often alienates Wikipedia's newcomers.

Leila Zia, the head of research at the Wikimedia Foundation, told me that her team was likewise working on tools that could help the encyclopedia by predicting, for example, whether a new article or edit would be overruled.

Or, she said, perhaps a contributor doesn't know how to use citations.

In that case, another tool would indicate that.

I asked whether it could help Wikipedia entries maintain a neutral point of view as they were writing.

Absolutely, she says.

For the moment, as the Wikipedia community debates rules and policy,

article submissions entirely written by LLMs are heavily discouraged on English language Wikipedia.

Still, there remains a kind of John Henry problem with AI.

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

The chatbots, unlike their human counterparts, have a formidable ability to churn out language like a steam-driven machine 24-7. I suspect the internet is going to be filled with crud just all over the place, Chris Albin told me. And with the AI models getting better at mimicking people's writing styles, it may be increasingly difficult to detect chatbot-written submissions. One Wikipedia editor, whose first name is Theo, sent me links in early June to show how he was in the midst of fending off a barrage of edits involving suspect citations formulated by AI, including one to an article about Lake Doxa in Greece. Often, I got the sense that Theo and other Wikipedians were worried that their human abilities to scrutinize new content and citations stretched to the limit already might soon be overwhelmed by an avalanche of AI-generated text. Certainly, new tools that were themselves AI would help. But even if the editors won in the short term, you had to wonder, wouldn't the machines win in the end? There are two sides to every story, but if you want to hire great talent for your business faster, there's just one way to do it. You need Indeed. They streamline hiring with powerful tools that find you matched candidates. Start hiring now with a \$75 sponsored job credit to upgrade your job post at [indeed.com-nyt](https://www.indeed.com-nyt). Offer good for a limited time. Just go to [indeed.com-nyt](https://www.indeed.com-nyt) and support the show by saying you heard about it on this podcast. [Indeed.com-nyt](https://www.indeed.com-nyt). Terms and conditions apply. Need to hire? You need Indeed. Three years ago, in anticipation of Wikipedia's 20th anniversary, Joseph Regal, a professor at Northeastern University, wrote a historical essay exploring how the death of the site had been predicted again and again. Wikipedia has nevertheless found ways to adapt and endure. Regal told me that the recent debates over AI recall for him the early days of Wikipedia

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

when its quality was unflatteringly compared to that of other encyclopedias. It served as a proxy in this larger culture war about information and knowledge and quality and authority and legitimacy. So I take a sort of similar model to thinking about chatGPT, which is going to improve. Just like Wikipedia is not perfect, it's not perfect. It's never going to be perfect. But what is the relative value given the other information that's out there? The future, as he saw it, would be a range of options for information, caveat emptor, including everything from chatGPT to Wikipedia to Reddit to TikTok. A dedicated plugin could, meanwhile, improve the chatbot's answers to questions about, for instance, health, weather, or history. At the moment, it goes against the grain to bet against AI. The big tech companies, wagering billions on the new technologies and largely undaunted by their shortcomings or risks, seem intent on forging ahead as fast as they can. Those dynamics would suggest that organizations like Wikipedia will be forced to adapt to the future that AI has begun to create, rather than exert influence over AI or mount an effective resistance to it. Yet many Wikipedians and academics I spoke with question any such assumption. Impressive as the chatbots may be, AI's apparent glide path to success may soon encounter a number of obstacles. These could be societal, as well as technical. The European Union's parliament is presently considering a new regulatory framework that, among other things, would force tech companies to label AI-generated content and to disclose more information about their AI training data. Congress is, meanwhile, considering several bills to regulate AI. Legal scrutiny may be coming, too. In one closely watched lawsuit, stability AI is being challenged for using pictures

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

from Getty Images without permission.

A California-class action suit accuses open AI of stealing the personal data of millions of people that has been scraped from the Internet.

While Wikipedia's licensing policy lets anyone tap its knowledge and text to reuse and remix it however they might like, it does have several conditions.

These include the requirements that users must share alike, meaning any information they do something with must subsequently be made readily available, and that users must give credit and attribution to Wikipedia contributors.

Mixing Wikipedia's corpus into a chatbot model that gives answers to queries without explaining the sourcing may thus violate Wikipedia's terms of use, two people in the open-source software community told me.

It is now a topic of conversation inside the Wikimedia community whether some legal recourse exists.

Data providers may be able to exert other kinds of leverage as well.

In April, Reddit announced that it would not make its corpus available for scraping by big tech companies without compensation.

It seems very unlikely that the Wikimedia Foundation could issue the same dictum and close its sites off, an action that Nicholas Vincent has called a data strike because its terms of service are more open.

But the Foundation could make arguments in the name of fairness and appeal to firms to pay for its API, just as Google does now.

It could further insist that chatbots give Wikipedia prominent attribution and offer citations in their answers,

something Selina Deckelman told me the Foundation is discussing with various firms.

Vincent says that AI companies would be foolhardy to try to build a global encyclopedia themselves with individual contractors.

Instead, he told me, there might be an intermediary stage here where Wikipedia says, hey, look at how important we've been to you.

Such an entreaty could be an effective reminder too

that the chatbots are made from us

without ingesting the growing millions of Wikipedia pages

or vacuuming up Reddit arguments about plot twists in *The Bear*.

New LLMs can't be adequately trained.

In fact, no one I spoke with in the tech community

seemed to know if it would even be possible to build a good AI model without Wikipedia.

It may require the equivalent of a death in the family

before the tech companies realize that they exist in a world of mutual dependency.

Already, according to the computer scientist working in the AI industry,

some technologists are concerned that new AIs are compromising the health of a website

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

for programmers called Stack Overflow, a popular platform that the models have been trained on to answer coding questions. The problem seems to have two distinct aspects. If those with coding inquiries can go to chatGPT for help, why go to Stack Overflow? In the meantime, if fewer people are consulting Stack Overflow for answers, why continue posting helpful suggestions or insights there? Even if conflicts like this don't impede the advance of AI, it might be stymied in other ways.

At the end of May, several AI researchers collaborated on a paper that examined whether new AI systems could be developed from knowledge generated by existing AI models rather than by human-generated databases. They discovered a systemic breakdown, a failure they called model collapse. The authors saw that using data from an AI to train new versions of AIs leads to chaos. Synthetic data, they wrote, ends up polluting the training set of the next generation of models. Being trained on polluted data, they then misperceive reality. The lesson here is that it will prove challenging to build new models from old models. And with Chatbots, Ilya Shemaïlov, an Oxford University researcher and the paper's primary author told me, the downward spiral looks similar. Without human data to train on, Shemaïlov said, your language model starts being completely oblivious to what you ask it to solve and it starts just talking in circles about whatever it wants, as if it went into this madman mode. Wouldn't a plug-in from, say, Wikipedia avert that problem, I asked. It could, Shemaïlov said.

But if in the future Wikipedia were to become clogged with articles generated by AI, the same cycle, essentially, the computer feeding on content it created itself would be perpetuated. Ultimately, the study concluded that the value of data from genuine human interactions will be increasingly valuable for future LLMs. At least for today's Wikipedians, that seems like encouraging news. Insofar as it suggests our new machines will need us, at least for a while, to keep them honest and functional and dependent on us. Ensuring that an AI system is doing what's in the best interests of humanity involves a theoretical concept known as alignment. Alignment is viewed as both an enormous challenge and an enormous priority for AI because a system out of sync with humans might create terrible damage. If AI ruins or compromises a mostly reliable system of free knowledge, it's difficult to see how that aligns with our best interests. One of the things that's really nice about having humans do the summarization is that you get some sort of basic level of alignment by default, Aaron Haffeker pointed out to me. And if you appreciate the editors of Wikipedia are human, they have human motivations and concerns, and that their motivations are providing high quality educational material to align with your needs, then you can essentially put trust in the system.

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

You can grasp the alignment argument better when you talk to people who devote their lives to the idea.

When I ask Jade, who has more than 24,000 edits to her credit, why she spends her free time, typically 10 to 20 hours a week editing Wikipedia, she said she believed in sharing knowledge.

Plus, I'm just a big nerd, she said.

We were speaking by Zoom late in the evening, and it was a conversation that had little resemblance to other long evenings of dialogue I'd had with chat GPT.

Some of Jade's work spoke to her personal interests in nature and birds, like an entry she wrote on the Vermillion Flycatcher, which got about 21,000 page views in the past 12 months.

She also told me she works regularly on the Wikipedia entry on the American Civil War, which had 4.84 million views over the same period.

Her goal was to continue to work toward completeness and greater accuracy in that Civil War article so that it achieves featured status on Wikipedia,

a rare recognition usually marked by a star of an article's quality

that is awarded by Wikipedia's editors to about 0.1% of English language entries.

My calculations in the past are, you know,

more than 10 million people read my work in a year, Jade said,

so it's an honor to have people reading all that.

We are going to have to create processes.

We are going to have to have hard conversations, she said, about the ethics of using AI to create Wikipedia articles.

When I asked her whether chatbots would soon eliminate her opportunities for volunteer work, she replied, I don't ever, maybe not never,

but certainly not in this century do I see robots fully replacing humans on Wikipedia.

I wasn't as sure.

The allure of a chatbot conversation, despite its factual shortcomings,

already seemed too irresistible and too enchanting to too many millions of people.

In fact, my own hours spent with chatGPT had chipped away at my own neutral point of view.

Not because the informational exchange was so rigorous and detailed, it wasn't,

but because the interaction was so captivating and effortless.

Nevertheless, Jade was resolute.

I'm an optimist, she said.

Heard of Instacart business?

It's a new way to get supplies and ingredients delivered in as fast as one hour, over 1200 retail brands.

Get delivery that moves as fast as you do.

Sign up for Instacart business and for a limited time,

get free delivery and 2% credit back for one year with an Instacart Plus trial.

Visit [instacart.com slash business](https://www.instacart.com/business).

Instacart Plus trial for new users while supplies last.

Plan renews at \$99 per year.

[Transcript] The Daily / The Sunday Read: 'Wikipedia's Moment of Truth'

\$250 credit back minimum.

Excludes alcohol.

Terms apply.