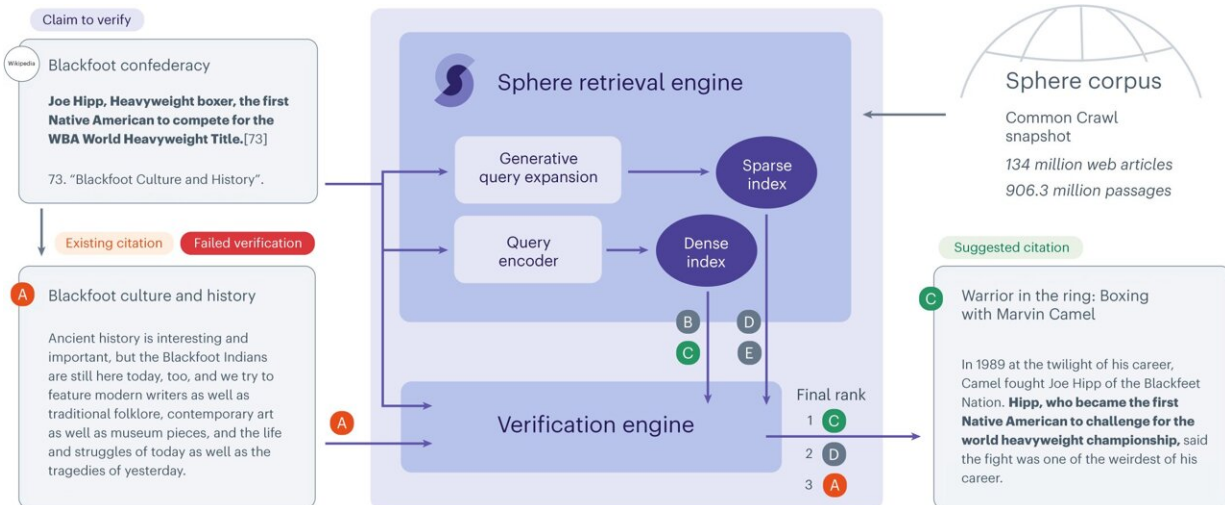


Study: AI can boost Wikipedia reliability

October 23 2023, by Peter Grad



The decision flow of SIDE from a claim on Wikipedia to a suggestion for a new citation is as follows: (1) the claim is sent to the Sphere retrieval engine, which produces a list of potential candidate documents from the Sphere corpus; (2) the verification engine ranks the candidate documents and the original citation with respect to the claim; (3) if the original citation is not ranked above the candidate documents, then a new citation from the retrieved candidates is suggested. Note that the score of the verification engine can be indicative of a potential failed verification, as the one reported in the example. Credit: *Nature Machine Intelligence* (2023). DOI: 10.1038/s42256-023-00726-1

There are enough differing opinions about the usefulness of Wikipedia to fill, well, an encyclopedia.

The novelist Nicholson Baker termed Wikipedia, a compendium of data its developers call a collection of all the knowledge in the world, "just an incredible thing. It is fact-encirclingly huge, and it is idiosyncratic, careful, messy, funny, shocking and full of simmering controversies—and it is free, and it is fast."

The writer Oscar Auliq-Ice declared Wikipedia "a revolutionary resource that has transformed the way people access and share information."

Some recognize the gigantic online resource—which as of this week contains over 6.7 million articles (in English) composed of more than 4.3 billion words—as a majestic but flawed undertaking.

"Wikipedia is like a flower bed, mostly beautiful with some ugly weeds," said environmental expert Steven Magee.

Noting the collective nature of Wikipedia, which is open to anyone volunteering information, Humorist Stephen Colbert suggested "Wikipedia is the first place I go to when I'm looking for knowledge ... or when I want to create some."

For some though, Wikipedia is a bitter pill. "I do not need wireless access to Wikipedia. I would prefer to stir-fry my own small intestines than to have continual access to a site where the entry for Klingon is longer than the entry for Latin," said Tara Brabazon, dean of graduate studies and professor of cultural studies at Charles Darwin University in Australia.

While generally regarded as a commendable, quick go-to source of information, users are always advised to do due diligence and not rely solely on one source of information. Rather, they should review other websites, explore article links and perhaps most important, check out

sources listed at the end of each Wikipedia entry.

Experts from around the world are regular contributors to Wikipedia, and most follow guidelines concerning neutrality and the use of reputable sources. The system generally works, though improvements can always be made.

This week, *Nature Machine Intelligence* [published](#) an article titled "Improving Wikipedia verifiability with AI," about a London-based AI company that seeks to bolster the [reliability](#) of Wikipedia's reference system. It does so by checking sources and identifying those that are accurate and those that are questionable, and then supplying its own recommendations.

Fabio Petroni, co-founder of Samaya AI, a knowledge-discovery platform, said, "The process of improving references can be tackled with the help of artificial intelligence powered by an information-retrieval system and a language model. Machines can help humans to find better citations, a task which requires understanding of language and mastery of online search."

His team trained their model on a huge dataset of Wikipedia entries, and then used it to review articles it had not scanned before. It analyzed sources and offered alternate reference sites, and its results were then examined by Wikipedia users.

When the AI system, called SIDE, classified Wikipedia sources as unverifiable and offered its own alternatives, users preferred SIDE's recommendations 70% of the time, the researchers found.

In about half of the cases, SIDE recommended the identical [sources](#) that Wikipedia had offered as its first reference.

"We demonstrate that existing technologies have reached a stage where they can effectively and pragmatically support Wikipedia users in verifying claims," Petroni said.

He said future research will focus on Wikipedia references beyond internet text, such as images, videos and paper publications.

"We hope that this work could be used in a broader context ... helping humans to check facts. More generally, we believe that this work could lead to more trustworthy information online," Petroni said.

More information: Fabio Petroni et al, Improving Wikipedia verifiability with AI, *Nature Machine Intelligence* (2023). [DOI: 10.1038/s42256-023-00726-1](https://doi.org/10.1038/s42256-023-00726-1)

© 2023 Science X Network

Citation: Study: AI can boost Wikipedia reliability (2023, October 23) retrieved 28 April 2024 from <https://techxplore.com/news/2023-10-ai-boost-wikipedia-reliability.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.