

Machine learning tool Quicksilver can find overlooked females

August 10 2018, by Nancy Owano



Wikipedia logo.

Primer's John Bohannon has been discovering people's work and contributions thanks to a machine learning system built at Primer. "It does this much as a human would, if a human could read 500 million news articles, 39 million scientific papers, all of Wikipedia, and then write 70,000 biographical summaries of [scientists](#)." The reason this is news is because you would probably never know about their contributions by frequenting Wikipedia, but he knows a way to fix that.

Wikipedia appears have a gender problem, for one. It is a matter of under-representation. And now the machine learning [system](#) at an AI startup has been showing how it could address the situation.

Primer is in the news. The Primer system was trained on scholarly journals. The gender gap-filling tool is called Quicksilver. It can spot many an overlooked female scientist with no presence at Wikipedia. Cory Doctorow in *Boing Boing* said 18% of Wikipedia's biographic entries were about women and the vast majority of Wikipedians were men.

The show-all process involved 30,000 Wikipedia entries to create a model that allowed it to identify characteristics making a scientist noteworthy for encyclopedic inclusion. Then, it mined the academic search-engine Semantic Scholar to identify 200,000 authors of [scientific papers](#).

Tom Simonite said in *Wired*: "Only 18 percent of its biographies are of women. Surveys estimate that between 84 and 90 percent of Wikipedia editors are [male](#)."

Actually, if you catch the story in *Wired*, the gender fix is part of the bigger story of Quicksilver looking for unsightly gaps.

In the bigger picture, blogged Bohannon, "Our aim is to help the open data research community build better tools for maintaining Wikipedia and Wikidata, starting with scientific [content](#).

(In addition, "Quicksilver doesn't just spot overlooked individuals and generate draft articles. It can also be used to maintain Wikipedia entries and [identify](#) when they haven't been updated for a while," said James Vincent in *The Verge*.)

So, what is the fix? Note that Primer is not about automated fixer-uppers. Said Simonite, "it doesn't plan to ever let Quicksilver autonomously add to the site." *Wired* quoted the CEO of Primer, Sean Gourley. "There are always humans in the loop." *Popular Science* said, "Quicksilver discovers scientists who should have Wikipedia articles about them and [writes](#) a first draft."

Their work continues. Bohannon said they have been quietly testing and improving Quicksilver for months. "Even before we finished the text generation component, Quicksilver was used in three English Wikipedia editathons for improving coverage of women of science. (Thank you to 500 Women Scientists for collaborating and inspiring us!)" He said they will describe their architecture in detail in future posts.

Meanwhile, wrote Simonite, "Wikipedia's notoriously punctilious community will likely keep a close [eye](#) on content generated with Quicksilver's help. One question is whether this tool aimed at fixing blind spots has any [blind spots](#) of its own."

© 2018 Tech Xplore

Citation: Machine learning tool Quicksilver can find overlooked females (2018, August 10)
retrieved 1 May 2024 from

<https://techxplore.com/news/2018-08-machine-tool-quicksilver-overlooked-females.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.