UC Irvine

UC Irvine Previously Published Works

Title

Node Position in a Group-Constructed Knowledge Network as an Indicator of Contributor Accuracy

Permalink

https://escholarship.org/uc/item/7cg9f4f2

Authors

Freedman, Hayden
Shaefer, David
van der Hoek, Andre
et al.

Publication Date

2022-07-12

Copyright Information

Peer reviewed

Node Position in a Group-Constructed Knowledge Network as an Indicator of Contributor Accuracy

Hayden Freedman (hfreedma@uci.edu)

David Schaefer (drschaef@uci.edu)

Andre van der Hoek (andre@ics.uci.edu)

Bill Tomlinson (wmt@uci.edu)

Group-constructed knowledge networks, such as Wikidata, are of increasing importance as corpora of knowledge. However, how to evaluate the performance of individual contributors to such knowledge networks is an open research question. We propose that using network measures to situate individual contributions relative to one another provides useful insight into their accuracy. In our study, 248 students in an undergraduate class about sustainability each generated 10-node networks, with nodes representing concepts from the course, and edges representing causal relationships between two connected concepts. Students were allowed to include any concept that appears in Wikipedia as a node, but were asked to use concepts that related to important course topics such as climate change. This common pool of nodes allowed us to link individual student content to create a bivariate consensus network with 734 nodes. For each student, we calculated several network metrics to summarize the position of their nodes in the consensus network. Each student's causal relations were then graded for accuracy and the student assigned a score. Our analysis found that several network-based summaries of student contributions (e.g., average betweenness) were significantly correlated with individual accuracy scores. This work demonstrates that the position of contributors' networks in a consensus network can indicate the quality of a contribution to the network. This finding could be useful for automatically detecting high or low quality contributors as part of efforts to improve the quality of Wikidata or other knowledge networks.