

UC Irvine

UC Irvine Previously Published Works

Title

Student-Generated Knowledge Graphs about Sustainability

Permalink

<https://escholarship.org/uc/item/3mw1t43b>

Authors

Freedman, Hayden

Van Der Hoek, Adriaan

Tomlinson, William

Publication Date

2021-06-21

Copyright Information

This work is made available under the terms of a Creative Commons Attribution-NonCommercial-NoDerivatives License, available at <https://creativecommons.org/licenses/by-nc-nd/4.0/>

Peer reviewed

H. Freedman, A. van der Hoek, B. Tomlinson. 2021. "Student-Generated Knowledge Graphs about Sustainability", International Symposium for Sustainable Systems and Technology 2021.

# Student-Generated Knowledge Graphs about Sustainability

Hayden Freedman ([hfreedma@uci.edu](mailto:hfreedma@uci.edu))

Andre van der Hoek ([andre@ics.uci.edu](mailto:andre@ics.uci.edu))

Bill Tomlinson ([wmt@uci.edu](mailto:wmt@uci.edu))

Knowledge graphs---interconnected networks of concepts and relationships---form the foundation for many computational efforts around the world. Google, Netflix, Facebook, and other major corporations maintain their own knowledge graphs. However, sustainability efforts are not always aligned with corporate goals; as such, organizational incentives may not be sufficient for the creation of knowledge graphs well-suited to support sustainability. A free and open platform called Wikidata run by the Wikimedia Foundation does exist, but it is currently sparsely populated with sustainability-related content.

We propose that there is a need for a free and open knowledge graph richly populated with sustainability knowledge, to support computational initiatives that seek to serve the public good. While there may be a lack of a corporate work force to generate such a sustainability knowledge graph, there are many thousands of students at universities around the world who are engaged in learning about sustainability. We see an opportunity for the work that these students do in their assignments to contribute to a sustainability knowledge graph, effectively crowdsourcing the effort.

As a first step, we conducted a study with 10 undergraduate students at the University of California, Irvine who recently completed an introductory sustainability-related course. These students were asked to individually create sustainability knowledge graphs. Participants were given 90 minutes to build a fully connected graph with at least 20 concepts and 19 relationships. We asked them to begin with the concept of "Sustainability" but gave no further instructions on what concepts to include. To ensure a controlled vocabulary of concepts, participants were limited to the use of Wikipedia article titles as possible concepts. We did not provide a controlled vocabulary for relationship labels, allowing participants to freely make associations.

After collecting students' individual graphs, we aggregated them into a single, integrated knowledge graph, which we then assessed for accuracy, relevance, and connectivity. We then compared the unified knowledge graph to the relevant subsection of Wikidata, assessing both

**H. Freedman, A. van der Hoek, B. Tomlinson. 2021. "Student-Generated Knowledge Graphs about Sustainability", International Symposium for Sustainable Systems and Technology 2021.**

how similar the students' work was to what is already known and what new relationships could potentially be contributed to Wikidata.

Results indicate that each participant was able to effectively create a knowledge graph with the required number of concepts and relationships. Participants collectively generated 172 unique concepts that spanned many different disciplines. Moreover, we found that the connectivity of the student-generated knowledge graph (270 relationships) was higher than the connectivity between the same concepts in Wikidata (86 relationships). However, the students used a relatively large set of relationship labels, employing 190 distinct labels for the 270 relationships. This indicates that limiting students to a controlled vocabulary of relationship labels may help students create stronger and more consistent associations that are better suited for incorporation into a larger knowledge resource. This study provides evidence that, when properly guided, undergraduate students may be able to contribute useful content to a shared data resource of sustainability knowledge. We envision the possibility of future software-supported curricula enabling students around the world to make many more contributions to shared public resources.