Proceedings of the Sixteenth International AAAI Conference on Web and Social Media (ICWSM 2022)

Improving Wikidata with Student-Generated Concept Maps

Hayden Freedman, André van der Hoek, Bill Tomlinson

Department of Informatics, University of California, Irvine
hfreedma@uci.edu, andre@ics.uci.edu, wmt@uci.edu

Abstract

Wikidata is a publicly available, crowdsourced knowledge base that contains interlinked concepts structured for use by intelligent systems. While Wikidata has experienced rapid growth, it is far from complete and faces challenges that prevent it from being used to its full potential. In this paper, we propose a novel method for improving Wikidata by engaging undergraduate students to contribute previously missing knowledge via concept mapping assignments. Rather than allow students to edit Wikidata directly, we describe a workflow in which knowledge is constructed by students and then reviewed by an expert. We present a case study in which we deployed a workflow in a large undergraduate course about sustainability, and find that it was able to contribute a substantial number of high quality statements that persisted in and contributed previously missing knowledge to Wikidata. This work provides a preliminary workflow for improving Wikidata based on classroom assignments, as well as recommendations for how future educational projects could continue to improve Wikidata or other public knowledge bases.

Introduction

Knowledge graphs are of growing interest in both computing research (Qi, Gao, and Wu 2017) and the computing industry (Noy et al. 2019). Wikidata, a site hosted by the Wikimedia Foundation, is perhaps the world's foremost example of a public, crowdsourced knowledge graph. It contains over 94 million concepts (e.g. *climate change*, *Ada Lovelace*, *Australia*) and over 1.2 billion links (Wikidata 2021a). Each link connects either two concepts or a concept and a value (e.g., "*climate change* has effect *sea level rise*", "*Ada Lovelace* instance of *human*", "*Australia* area *7,692,024 square kilometers*"). In spite of its rapid growth over the past decade, Wikidata faces a number of unique challenges (Spitz et al. 2016). In particular, Wikidata is not complete and contains gaps in knowledge which often take the form of missing links. In spite of significant and ongoing research in the fields of link prediction and information extraction (Cao et al. 2021; Ali et al. 2019), missing links remain a problem in Wikidata (Shenoy et al. 2021).

Filling in knowledge gaps in Wikidata is important as its data is used in several high impact areas of the Internet, including Wikipedia infoboxes (Vrandečić 2013) and Google Search knowledge panels (Schwartz 2014), and as a knowledge corpus for artificial intelligence agents (Acosta et al. 2019).

Given the importance of Wikidata as a public resource, as well as the ongoing efforts to improve the quality of its data, in this paper we address the following research question:

Can students generate high-quality knowledge that can be used to fill in missing links in Wikidata as part of their course assignments?

In this paper, we discuss a novel means of improving Wikidata via concept mapping assignments deployed in classrooms. We present results from this method being deployed in an undergraduate sustainability course at a large US university in Spring 2021. Students generated concept maps based on their understanding of course material; throughout the quarter they refined their maps by pruning and adding statements. An expert reviewed each statement in each final concept map, determined its appropriateness for Wikidata, and uploaded the selected statements to Wikidata. The benefit of this workflow to the expert is that it transforms their task from one of generation of content to one of evaluation of content. The difference in cognitive load between having an expert generate a body of content versus evaluating the quality of that content is akin to the difference between recognition and recall (with recognition being faster and easier than recall) (Budui 2014). Therefore, workflows such as this one may lower the amount of expert work needed to populate Wikidata.

Research has separately demonstrated the value of both crowdsourcing domain-specific knowledge (Hammon and Hippner 2012; Chilton 2009; Griffith et al. 2017) and concept mapping activities as an educational tool (Novak 1990; Kinchin 2014; Plotnick 1997). These two concepts have some symmetry, since concept mapping assignments are inherently a form of structured knowledge production. In spite of this, no prior work exists that uses the domain-specific knowledge students create via concept mapping as a means of improving public knowledge representation. This paper addresses this gap and thereby makes the following research contributions: (1) a preliminary methodology for generating high-quality structured knowledge via concept mapping assignments, (2) the addition of several hundred class and instance-level statements to an underrepresented region of Wikidata; and (3) an analysis of the contributed knowledge. This paper also includes recommendations for how future assignments could help students best engage with and contribute to high-quality, public knowledge bases via their coursework.

Background and Related Work

Wikidata as Social Media

Public knowledge bases are of growing importance in the digital world (Noy et al. 2019). Wikidata has roughly 23,000 active users, who manually edit concepts and links, oversee bots that automatically populate Wikidata from external databases, and engage in community discussion and activity. Each month, nearly 20 million edits to Wikidata are made manually by users and by automated bots, and more than 2,000 new users join the site, making it one of the most active Wikimedia projects (Wikimedia 2021; Wikidata 2021b). Users can interact directly with Wikidata by viewing specific concepts and their outgoing links via the web interface, adding and removing concepts or links, or using the Wikidata Query Service (MediaWiki 2021) to launch programmatic SPARQL queries to retrieve specific data.

Since its inception in 2012, one of Wikidata's main uses has been to provide up-to-date, machine-readable data to AI services such as the Google Knowledge Graph, Wolfram Alpha, and IBM Watson (Vrandečić and Krötzsch 2014). In 2014, Google committed to transferring all content from Freebase, a Google-owned knowledge graph, to Wikidata (Pellissier Tanon et al. 2016), indicating their intention to continue relying on Wikidata to help populate Google Search results. However, Google does little to support Wikidata other than exploit the knowledge it contains, and may even reduce traffic to Wikidata by displaying knowledge in its own search panels without crediting Wikidata (McMahon, Johnson, and Hecht 2017).

A wide variety of domains and applications have used Wikidata for various purposes, including constructing biomedical domain ontologies to assist clinical decision support systems (Waagmeester et al. 2020); enabling the search of multilingual lexemes (Nielsen 2019); automatically generating scholarly profiles for researchers, organizations, and institutions (Nielsen, Mietchen, and Willighagen 2017); detecting fake news (Dun et al. 2021); and NLP tasks such as named entity recognition (Satyapanich, Ferraro, and Finin 2020). Wikidata also provides the data behind many of Wikipedia's infoboxes (Lemus-Rojas and Pintscher 2017), ensuring that its data is readily available to users responsible for Wikipedia's 18 billion monthly page views (Anderson, Hitlin, and Atkinson 2021).

Concept Mapping

Concept mapping is an educational framework that has been shown to be an effective means of facilitating student understanding of the complexity and interrelatedness of scientific concepts (Tan, Erdimez, and Zimmerman 2017; Novak, Gowin, and Kahle 1984; Maker and Zimmerman 2020). Concept maps are visual representations of domain concepts and the links between them. For example, a simple concept map about the Solar System might include concepts for planets, moons, and the Sun, and utilize the link label "revolves around" to connect the concepts in a logical manner. Concept maps can be used both to measure student understanding and facilitate it, and have been demonstrated to be effective educational tools in environments ranging from elementary school to undergraduate courses, and in scientific disciplines as varied as computer science, medicine, and psychology (Cui and Yu 2019; Dos Santos et al. 2017; Daley, Durning, and Torre 2016; Siew 2019).

As a highly interdisciplinary and complex domain with many dependencies and links among concepts, sustainability is well suited to be taught using concept mapping activities (Segalàs, Ferrer-Balas, and Mulder 2008, 2010; Walshe 2008; Bielefeldt 2016; Shallcross 2016). Sustainability concept mapping activities have been shown to increase activity in regions of students' brains linked to systems thinking compared to simply listing concepts (Hu et al. 2019). Maher and Fisher (Maher and Fisher 2012) discuss how AI-supported social construction of knowledge can be applied to build a sustainability concept map, which could then be used as the building blocks of further knowledge construction. The authors felt that a sustainability concept map would be valuable compared to other possible topics because sustainability "has many and possibly conflicting views", making it also interesting for the case study described here because of the wide variety of possible concepts and links that students could construct related to this topic.

Crowdsourcing

Crowdsourcing is a growing subfield of computing, with implications for many other areas of computing and for other fields as well. Crowdsourcing involves "a large group of people (a crowd) work[ing] on solving a problem, providing data or contributing to a common goal" (Jäger et al. 2019). Over the past decade, crowdsourcing has seen a rapid growth in interest from the scholarly community (Doan, Ramakrishnan, and Halevy 2011; Silberman et al. 2018; Salehi et al. 2017). Broadly, crowdsourcing is often used for social good, such as efforts to understand how people make moral decisions (Awad et al. 2020), to engage in citizen science (NOAA 2020), and to pursue collective problem solving (Duhaime, Olson, and Malone 2015). Specific applications include collecting data on human mobility (Consonni et al. 2021), annotating song lyrics (Lim and Benson 2020), and characterizing abusive online behavior (Founta et al. 2018). Wikipedia remains one of the Internet's most famous examples of crowdsourced labor working towards a common goal (Libert, Spector, and Tapscott 2007).

Previous work has explored how crowdsourcing can help aid sustainability initiatives. Several projects have used crowdsourcing to collect online public sentiment data related to sustainability and climate change (Abbar et al. 2016; Diakopoulos et al. 2014), while others have deployed gamified apps that seek to educate users on sustainable practices while collecting environmental data (Piccolo et al. 2016). Other research proposes a framework for using crowdsourced environmental data to normalize environmentally-friendly behavior (Massung and Preist 2013), discusses the creation of an application that brings communities together to discuss and work on local sustainability issues (Lee et al. 2013), and studies the possibility of community-based air quality monitoring (Aoki et al. 2017).

Evidence suggests that obtaining high quality data from a distributed workforce is not trivial (Mason and Watts 2009; Shaw, Horton, and Chen 2011; Kazai, Kamps, and Milic-Frayling 2013; Mitra, Hutto, and Gilbert 2015; Hung et al. 2015). The task of populating Wikidata involves sufficient complexity that it is nearly impossible to avoid errors or inconsistencies. The research described here explores ways in which students may have a pedagogically sound learning experience while simultaneously improving the quality of content in a crowdsourced knowledge base.

Case Study

Deployment of Sustainability Concept Map Assignments

In the research underlying this paper, we used a concept mapping assignment in a large undergraduate sustainability course to simultaneously serve the pedagogical goals of the course and enable novel contributions to improve Wikidata. Despite the abundance of content on Wikidata, the robust population of editors, and a wide variety of practical use cases, there is still much room for improvement in Wikidata. The case study described here applies a workflow to the content base of Wikidata in the sustainability domain, but also considers broader applicability to Wikidata as a whole.

In this study, we implemented three concept mapping assignments in a large undergraduate course on sustainability and computing. Two of the authors of this paper were the instructor and a teaching assistant for this offering of the course. Before completing any concept mapping activities, students in the course were presented with a study information sheet to read and asked whether they would like to participate in the research study. A total of 248 students were enrolled in the course, of whom 139 consented to participate, were over the age of 18, and submitted the final concept mapping assignment. None of the authors had access to data on which students consented until after the course was over and final grades were submitted. All procedures were approved by the university's Institutional Review Board (IRB).

As the first step, students were asked to familiarize themselves with the freely available concept mapping software CmapTools (Cañas et al. 2004), and then to use this tool to create their own concept maps related to the course material (readings, lectures, etc.). Students were provided with a set of 20 link labels that they were allowed to use in their maps. This list was constructed by two of the authors of this paper (the instructor and teaching assistant for the course) reviewing the full list of link labels currently permitted in Wikidata (known in Wikidata as properties), and developing their own list of link labels that would be most relevant to the topics of the course. For example, the link label "causes" was selected because of the importance of documenting causal links in environmental phenomena. The majority of the selected link labels had obvious mappings to Wikidata properties; however, in a few cases the research team decided to add link labels that did not have obvious Wikidata mappings because they felt these link labels in particular would provide students with more flexibility with which to craft their concept maps, and were therefore pedagogically important. One such example is "inhibits", which does not exist in Wikidata but was perceived as important in order to allow students to make statements about interventions that have a dampening effect on climate change or other negative environmental phenomena.

Students were also required to represent concepts using Wikipedia articles (which directly map to Wikidata entries by definition of how Wikidata is constructed). This restriction was imposed for several reasons: to ensure that students did not refer to the same concept by different names, to ensure that the concepts students referenced would exist in Wikidata, and to provide canonical definitions in the form of Wikipedia articles in the case that the definition of a concept was ambiguous. Due to this restriction, all of the concepts included by students in their maps were already present in Wikidata; however the majority of the links that students built between the concepts did not yet exist in Wikidata.

In the first assignment, each student built a concept map associating the concepts of *sustainability* and *technology*, the two central themes of the course. We required that these two concepts be present to ensure that students had a starting point for their work and stayed on topic; students were also required to include at least 8 other concepts and at least 9 directed, labeled links. Additionally, students were not allowed to have disconnected components in their maps, so that each included concept was directly or indirectly associated with one of the main themes of the course. The vast majority of students were successful in producing concept maps as specified. The 139 students collectively produced a total of 1,859 links in the first assignment.

In the second and third assignments, students were asked to review the concept maps that they created in the prior assignment, and make at least 10 changes. Four types of changes to the maps were allowed: removing a link between existing concepts, changing the label of a link between existing concepts, adding a new link between concepts, or adding a new concept and connecting it to the rest of the network via a link. Most changes that students made in these assignments involved the adding of new concepts and links rather than modifying existing links. Students in the study collectively added 1,514 links from the first assignment to the second, and 1,472 links from the second assignment to the third. Figure 1 shows part of one student's final concept map in the CmapTools software. The student whose work is shown here developed a relatively complex concept map, having included many of the important course concepts and a thorough network of labeled links connecting the concepts. Beyond the concepts discussed in the course, this student also included the concepts of *bacteria* and *antibiotic*, which were not related to the course and may relate more to the student's own background or interests.

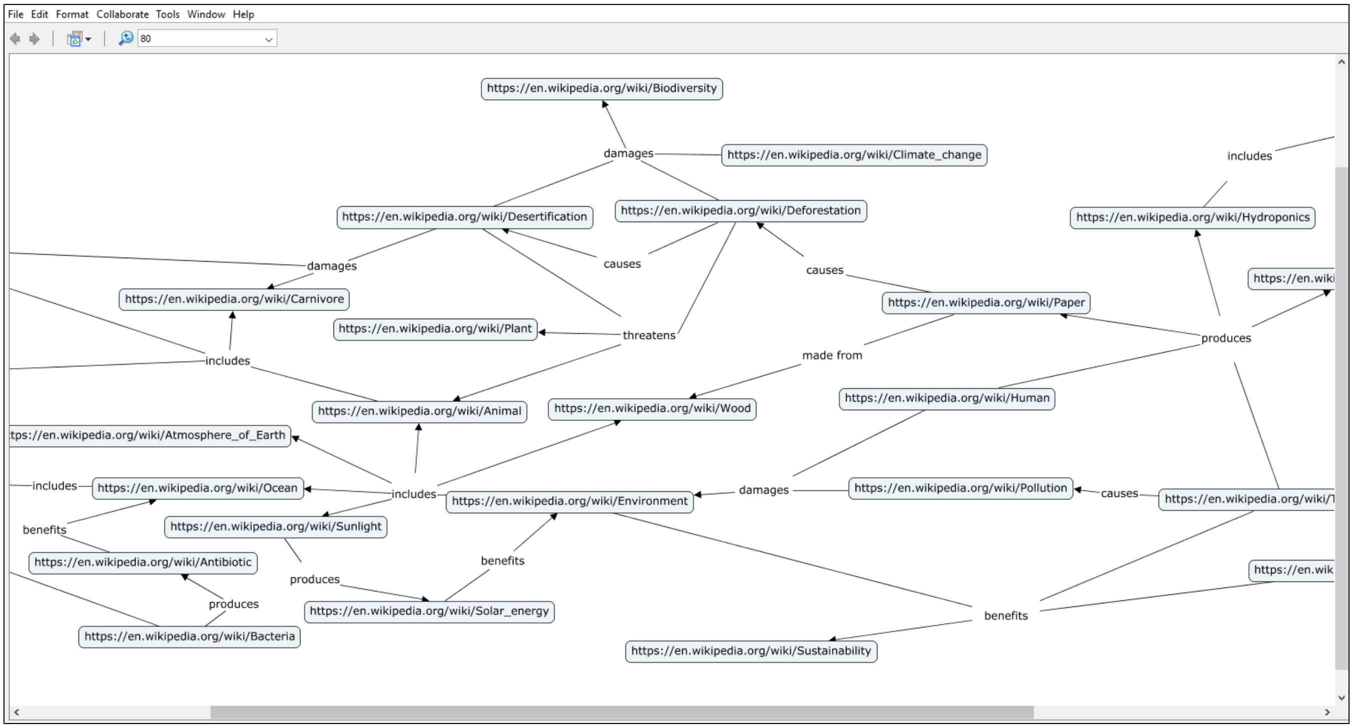

Figure 1: CmapTools environment for student concept map construction

Assigned Category	**Num. Statements (%)**
Total Inappropriate for Wikidata	2,917 (70.7%)
Total Appropriate for Wikidata	1,208 (29.3%)
- *Appropriate but Incompatible with Wikidata*	*622 (50.9%)*
- *Appropriate and Added to Wikidata*	*402 (33.9%)*
– Added as is	280 (68.6%)
– Added but modified by expert	122 (31.5%)
- *Appropriate but already present in Wikidata*	*137 (11.3%)*
- *Duplicate of other Appropriate Statement*	*47 (3.9%)*

Table 1: Results of Student Concept Map Review

Expert Review

After the students completed their third and final iteration of the concept mapping activities, the first author reviewed each statement from the concept maps generated by the 139 students in the study, a total of 4,125 unique statements. As a PhD student with a long-standing interest in sustainability, the first author had developed expertise on a range of sustainability-related topics. In addition, as a teaching assistant for the course, the first author was intimately familiar with the topics being studied. The first author judged each statement to be in one of three categories: Appropriate for Wikidata, Inappropriate for Wikidata, and Incompatible with Wikidata. Statements that were deemed Appropriate for Wikidata were factually correct and not overly general (using the first author's previous experience with Wikidata regarding generality)[1]. Statements marked as Incompatible with Wikidata were considered of sufficient quality but violated a Wikidata property constraint. As an example, the statement *waste* physically interacts with *ecosystem* was not allowed because the Wikidata property "physically interacts with" requires a subject that is classified as a *physical object* and the concept of *waste* is not a direct or indirect subclass of *physical object*.

Each statement in the "Appropriate" category was then added to Wikidata by the first author via the Wikidata web interface. In some cases during this process, the original stu-

[1]While Wikidata is at its core a corpus of facts, it is also a hierarchy of classes. Introducing accurate but overly general knowledge subverts the subclass hierarchy and likely reduces, rather than improves, the quality of Wikidata. The first author used their previous experience with Wikidata as well as with ontological modeling to determine which statements were too general for inclusion, or needed to be modified from their original form for inclusion. We discuss the discretionary role of the expert in the workflow, as well as how that role might be reduced in future iterations, in the Discussion section.

dent statement was made more specific or more general in order to make the statement more compatible with Wikidata's class hierarchy. For example, the statement *tropical rainforest* includes *understory* is correct, but since all forests contain an understory, the first author generalized the statement to *forest* has part *understory*, and then ensured that *tropical rainforest* was listed as a direct or indirect subclass of *forest*.

Revisions by the first author, as described above, were made for about 32% of the statements that were added to Wikidata. These revisions were motivated by a general philosophical aversion of the research team to add any links that could potentially reduce the quality of Wikidata, and an understanding that this workflow is a prototype rather than a full-fledged crowdsourcing workflow. We view the statements that were changed by the first author as an opportunity to observe patterns in the ways that student statements are accurate but imprecise, which can help reduce the work required between student submission and Wikidata contribution in future iterations of this research. We return to this topic in the Discussion section.

The results of the expert review of the student concept maps are shown in Table 1. The majority of the student-generated statements were deemed inappropriate for Wikidata. Types of statements in this category include statements that were factually incorrect or unsubstantiated (*ecological stability* has purpose *politics*), were considered too general to be useful (*Earth* studied by *scientist*), or potentially true in some cases but with obvious counterexamples (*environmental technology* has purpose *renewable energy*, since *renewable energy* is only one of many possible purposes of *environmental technology*).

The next largest group contains statements that were considered of appropriate quality for Wikidata but were unable to be added, because they would have introduced a constraint violation based on the Wikidata property they use. Wikidata does allow users to add statements that violate its constraints, and in some cases Wikidata constraints are not appropriate and should be changed, but knowing this would require further research on the constraints of each Wikidata property in question, which was beyond the scope of the current study.

Analysis of Sustainability Knowledge Contributed to Wikidata

Quantity of Knowledge The roughly 10% of student statements that were deemed both appropriate for and compatible with Wikidata were added by the first author over the course of two weeks in mid-August 2021, via the Wikidata web interface. By mid-December 2021, each statement had been available for Wikidata moderators to review for about 4 months. The vast majority of the approved statements remained in Wikidata in their original form close to the time of paper submission, as discussed later in this section. Figure 2 shows a screenshot of Wikidata's page on *plastic pollution*, with the red boxes showing the contributions from the workflow described in this paper.

Figures 3 and 4 are visualizations of the contributions at different scales. In each of the figures, Wikidata concepts are

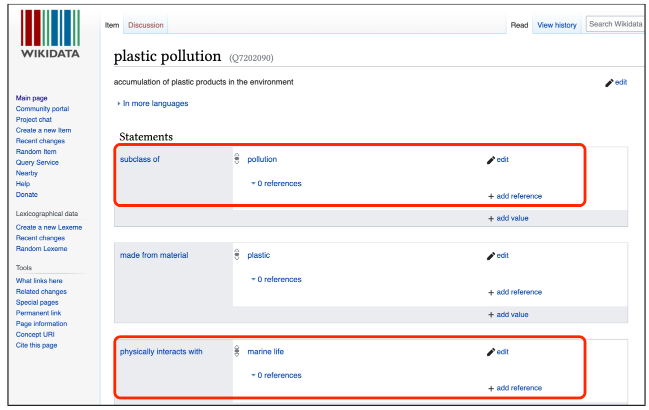

Figure 2: The Wikidata page for *plastic pollution*, including links that were generated by students, evaluated by an expert, and uploaded to Wikidata (highlighted in red). A total of 402 statements, connecting 412 concepts, were uploaded as part of this research.

shown as black circles. Pre-existing connections are shown as blue links between pairs of concepts, and contributions added via this project are shown as red links. Figure 3 shows all of the links between the 412 concepts to which the workflow added at least one link, before and after the contributions were made. Figure 4 shows all of the links between the concepts directly connected to the concepts of *sustainability*, *pollution*, and *urban agriculture*, before and after the contributions were added. In total, the workflow added links between 412 distinct Wikidata concepts, increasing the connectivity among this set of concepts by 57.4%.

Our findings indicate that the student-generated statements that were added to Wikidata were able to contribute substantially to Wikidata's sustainability region, helping to address the issue of missing links in this domain. In addition, the workflow contributed a substantial number of the connections between several of the Wikidata concepts that students linked more frequently in their own concept maps. Table 2 shows the 10 Wikidata concepts that gained the most connections to other concepts due to these contributions, and the percent change in the number of incoming and outgoing links to other concepts for each. The percent change in links varies widely based on the number of links that were already present for each of these concepts: while the contributions form a relatively small percentage of the links to or from *climate change* and *greenhouse gas*, they had a much larger relative impact on concepts with fewer pre-existing links such as *sustainable energy* and *pollution*. As expected, all of the top concepts are important elements of the course that students learned about throughout the quarter.

The contributions also filled in a substantial number of links for many important domain-specific concepts. They comprise 5 of the 8 listed effects of *deforestation*, 8 of the 20 effects of *climate change*, and all 6 of the concepts that have an objective or goal of *sustainability*. The workflow also contributed all 3 listed effects of *electronic waste* (*environmental degradation*, *global waste trade*, and *pollution*), the

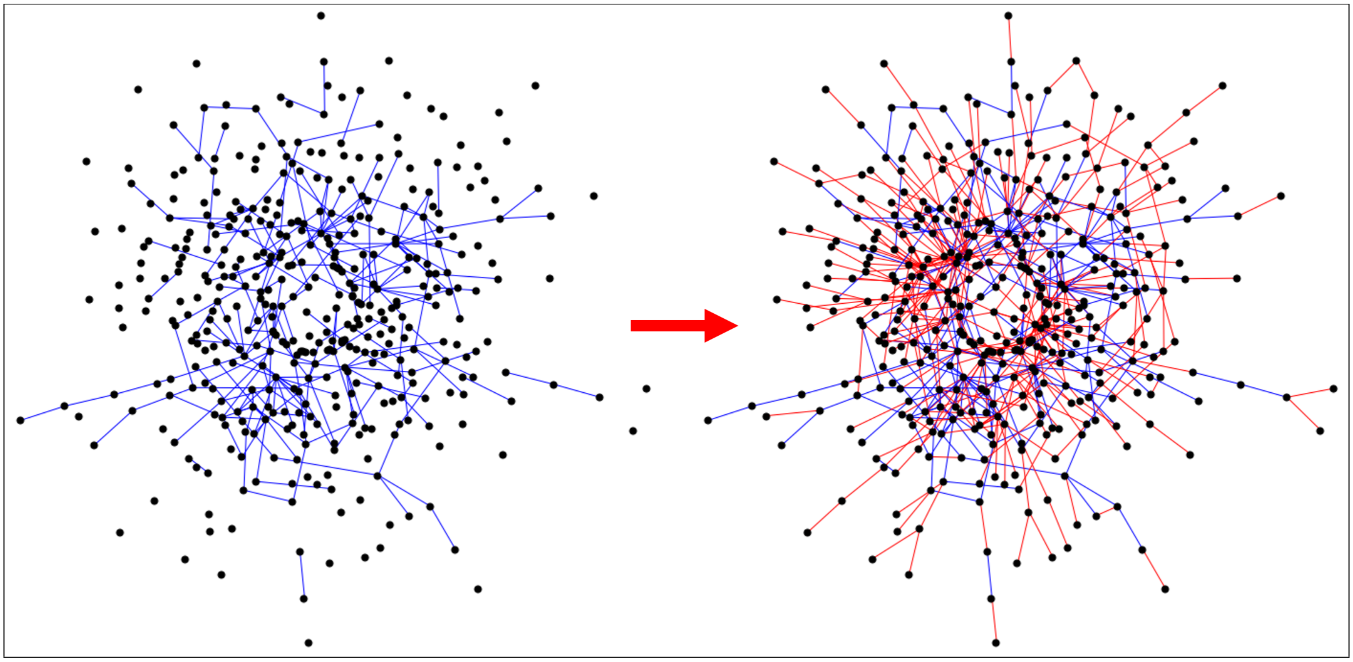

Figure 3: All Wikidata concepts to which the workflow contributed at least one link. The left network displays the concepts to which statements were contributed, and the pre-existing links among them. The right network includes the same concepts, with both pre-existing links (blue) and those added by the workflow (red). This figure demonstrates that the contributions via the research described here led to a non-trivial enhancement of this region of Wikidata.

Concept	Added Links	% Change in Links
sustainability	25	3.87%
climate change	19	0.09%
air pollution	16	0.12%
pollution	15	12.93%
sustainable energy	13	39.39%
natural environment	13	4.42%
deforestation	12	1.13%
aquaponics	10	125.00%
fossil fuel	10	6.62%
greenhouse gas	9	0.30%

Table 2: Wikidata Concepts with Most Links Added

only 2 listed effects of *habitat destruction* (*biodiversity loss* and *extinction*), and all 3 by-products of *coal-fired power stations* (*sulfur dioxide*, *carbon dioxide*, and *fly ash*).

Moderation of Knowledge by Wikidata Editors The statements that the workflow contributed to Wikidata as part of this research project have overwhelmingly remained live on the site. Of the 402 statements we uploaded, 392 (97.5%) were still present on the live Wikidata site after four months, 5 (1.2%) had been modified, and 5 (1.2%) had been removed. As an example of a modification, five days after the statement *photovoltaic system* part of *low-energy house* was contributed, a Wikidata moderator generalized the statement to be *photovoltaic system* part of *energy-efficient building*. Nearly all of this moderator activity happened within 2.5 weeks of the initial contribution, with only one additional edit being made after the first 2.5 weeks, presumably because recent additions to Wikidata are of particular interest for Wikidata moderators.

Since statements have no trace of having been reviewed within Wikidata unless they are removed or changed, we were unable to determine the proportion of added statements that were reviewed by Wikidata editors. This fact leaves open the possibility that some of the added statements that remain in Wikidata do so not because of their merit but because no moderator has looked at them. We believe that the reviewing on Wikidata is as yet significantly less thorough than that of its sister project Wikipedia. Nevertheless, 89% of pages to which the workflow contributed were edited at least once over the four months since the statements were added, indicating that other human moderators or bots were actively working to improve this set of pages. The fact that some of the statements were revised/removed demonstrates that there was some degree of oversight, and the fact that nearly all of the statements remained in place after the four-month period provides some implicit support for the appropriateness of those statements.

Recommendations for the Design of Knowledge-Generating Class Assignments

We view the mission of this research as important, as it has the potential to make substantial improvements to a widely used, crowdsourced knowledge base while helping students learn important course concepts via visualized knowledge

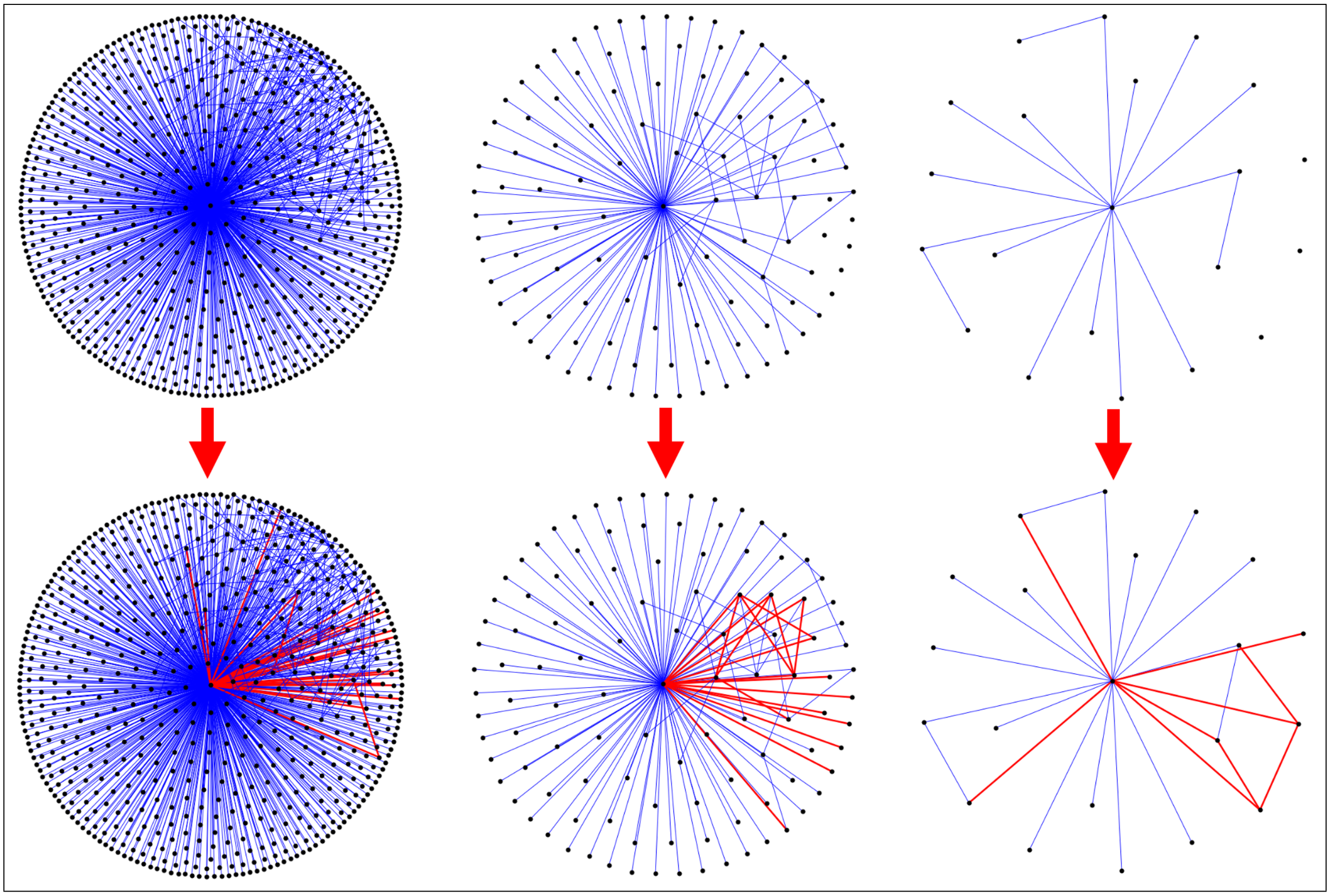

Figure 4: This figure shows pairs of charts for three concepts in Wikidata, with the top row showing pre-existing links between each concept and related concepts, and the bottom row showing the pre-existing links plus the links added by the workflow. The two networks at left show *sustainability*, the two center networks show *pollution*, and the two at right show *urban agriculture*. This figure shows that, while concepts had a range of densities of pre-existing links, the workflow was able to add novel links to these concepts. (Note: the majority of pre-existing links to the *sustainability* concept are academic papers on that topic.)

construction. The case study was promising, as we were able to engage students in creating a substantial body of structured knowledge within the sustainability domain, a subset of which has been added to and persisted in a public knowledge resource. Based on the case study, we make the following recommendations for future workflows attempting to use course assignments to improve the quality of public knowledge bases:

Ensure Compatible Form Factor and Vocabulary Future workflows should consider the shape of data that will be outputted by student assignments and that which is required by the target knowledge base. Concept maps worked well in the case study because they not only encourage students to think about the interconnectedness of concepts, but they also produce concept networks that are shaped similarly to Wikidata's graph structure. Along the same lines, it was crucial that students started off working with a common set of concepts and links that were also present in the target knowledge base. Due to these similarities, it was trivial to slot individual statements from the concept maps into Wikidata where appropriate, which would not have been the case if the student output had been in a different format. Additionally, a common set of concepts prevents students using different names to refer to the same thing, which would then become an entity resolution task.

Provide Students Opportunities to Self-Validate In the case study, students submitted an initial concept map and then two subsequent maps in which they made improvements on the previous iteration. We analyzed the statements that were removed between each iteration and found that the removed statements were of substantially lower quality than the statements that persisted, indicating that the iterative process was effective in producing higher quality content than a single concept map assignment would have been. Furthermore, we observed that statements made by multiple students were substantially more likely to be accurate than statements made by only a single student. These findings are in line with past research on crowdsourcing systems that use multiple assignment of tasks and majority voting to increase reliability (Karger, Oh, and Shah 2014).

We feel there is room for educators to be creative with student validation; having students peer review each other's assignments or do group-based editing of a merged concept map could be interesting ways to further improve the quality of knowledge. Such endeavors allow students to be more involved in the process of creating high-quality knowledge while engaging with course concepts. In the case study, a relatively small percentage (10%) of student-generated knowledge was actually added to Wikidata; more intensive student validation might help to raise this number in subsequent workflows.

Target Assignments Towards Improving Specific Aspects of the Knowledge Base In our review of previous research, we identified several areas where Wikidata could be improved, and then chose to focus specifically on the issue of missing links. As such, we tailored the course assignments to generate knowledge that could be used to fill in these gaps. In this section, we briefly discuss two other areas in which Wikidata could be improved, and propose assignment sequences that could help students engage with these improvements.

Inaccurate Information: In this assignment sequence, students are presented with subgraphs of Wikidata related to a topic of interest and asked to make improvements. In this case, rather than building concept maps from scratch, students engage with the knowledge already in Wikidata, and determine which statements are accurate and which require modification or deletion. Students propose their own modifications or additions to the subgraph they receive, and their recommendations are then reviewed by an expert.

Inconsistent Assignment of Classes and Instances: Unlike many knowledge graphs, Wikidata does not have a defined hierarchy of classes, leading to inconsistencies (Spitz et al. 2016). Furthermore, many instances are not assigned to an appropriate class, which recently hindered efforts to detect inaccurate information about Covid-19 in Wikidata (Turki et al. 2020). To help address these issues, this assignment sequence involves presenting students with a concept map showing a subset of the Wikidata class hierarchy, as well as instances assigned to each of the classes. Students are asked to identify inconsistencies in the class hierarchy as well as instances that are incorrectly assigned to classes. This assignment could be particularly of interest to introductory computer science students as a means of teaching the distinction between instances and classes, which is an important concept in object-oriented programming.

Discussion and Future Work

Concept mapping activities are often implemented in part to elicit students' cognitive structures as a means of approximating what they do and do not know about a given topic. A useful (and to this point, unrealized) side effect of these activities is that the student concept mapping data inherently includes structured, domain-specific knowledge. The research we describe in this paper supports the creation of novel workflows that allow students to contribute to public knowledge resources such as Wikidata through their assignments. Although the case study focuses specifically on the sustainability domain, nothing about this methodology is specifically tied to it, and we feel this approach would be generalizable to other domains and topics.

A large challenge with the workflow, related to the low overall quality of the student-generated data, involves most effectively using expert time. We do not expect that any workflow without an expert-in-the-loop would be philosophically reasonable, due to students' lack of expertise and the difficulty for an AI agent working without expert oversight to assess whether a given statement is appropriate for Wikidata. This work operates on the basis that evaluation of content is less time consuming for experts than generation of content, and is therefore more efficient for the system as a whole (since expert time is typically scarcer than student time). However, future work might seek to investigate how expert time could be further reduced in subsequent workflows. For instance, a modification of the probabilistic model proposed by (Hung et al. 2015) to knowledge graphs could serve to substantially reduce expert work by providing an order of student-generated statements ranked by their likelihood of appropriateness for Wikidata. Future work should also investigate the time commitment required to contribute to the workflow, from the perspective of both students and experts.

A further challenge in working with the student-generated data was that a portion of the data required modification by the expert before being added to Wikidata. Most of these modifications were due to students using Wikidata classes that were either too specific or too general (as discussed in the Expert Review section). This problem placed an additional burden on the expert, as they not only had to judge the accuracy of each statement but also analyze the Wikidata class hierarchy to ensure that each statement was being added at the appropriate level of the hierarchy. Given that the students were generating knowledge from scratch without knowledge of the current shape of Wikidata, it is unsurprising that we encountered this issue. To further reduce the burden on the expert, we will investigate the effects of integrating the Wikidata class hierarchy into the concept mapping software, such that students are prompted to consider all direct subclasses and superclasses each time they add a new concept to their concept map. As noted in the previous section, the Wikidata class hierarchy is not perfect; therefore such an integration could simultaneously be aided by student efforts to improve the class hierarchy.

As an additional point, while student experience is not the focus of this paper, we are nevertheless encouraged that students in the sustainability course scored the concept mapping activities as the most engaging assignments in the course on an end-of-course evaluation. There has been significant research showing that projects with a social impact can be motivating for students, and especially for female students and students in other groups underrepresented in computing (Margolis et al. 2017; Margolis and Fisher 2002). With 220 million tertiary students in the world at any given time, many of whom are doing hundreds or even thousands of hours of work per year on their studies (WorldBank 2021), harnessing even a small subset of these billions of hours of effort could be hugely valuable to shared knowledge resources. We also recognize that students learn better from "authentic assignments" that mirror real-world tasks (Steiner 2016). We strive to create a framework that not only improves a public knowledge base but also forms a valuable piece of an education curriculum.

Finally, an additional area of future study is to examine differences in student performance across groups, such as academic discipline. When we grouped students by school and evaluated the performance of these groups in terms of the quality of their contributions, we found some substantial differences. For example, students in the School of Humanities contributed high-quality statements at a rate of approximately 4 times that of students in the school of Biological Sciences, and 5 times that of students in Social Ecology. These observations are difficult to explain offhand, and would be an interesting source of further research.

Conclusion

Wikidata is an important public knowledge resource; as such, the data it contains should be as complete and as accurate as possible. To explore a new way of making long-lasting contributions to Wikidata, we presented a case study in which undergraduates in a sustainability course created concept maps that were vetted by a member of the research team and contributed to Wikidata. In this process, the contributions constituted a substantial increase in the number of links among key concepts such as *sustainability*, *climate change*, and *air pollution*. We provided recommendations for how instructors might create similar assignments to further involve students in the improvement of Wikidata. We envision this work as a first step toward an array of processes by which students across many different disciplines may help improve Wikidata or other public knowledge bases as part of their educational curricula.

Ethics Statement

The authors have no conflicts of interests to declare, and believe that the work described in this paper meets the standards of the AAAI Code of Professional Ethics and Conduct. We believe that the enrichment of Wikidata with sustainability knowledge is both socially responsible and broadly accessible. We recognize that crowdsourcing efforts may lead to the introduction of bias in datasets (Ghai et al. 2020); in future work, it will be relevant to determine in what ways bias introduced by students is greater than, less than, or different from bias introduced by other crowds. And, while AI systems involve many ethical challenges (Coeckelbergh 2020), this work seeks to produce a more robust knowledge corpus that may enable AI systems to contribute more effectively to the transition to sustainability. By doing so, it may contribute to the well-being of marginalized groups that are most likely to be affected by planetary issues such as climate change.

References

Abbar, S.; Zanouda, T.; Berti-Equille, L.; and Borge-Holthoefer, J. 2016. Using Twitter to Understand Public Interest in Climate Change: The case of Qatar. In *Tenth International AAAI Conference on Web and Social Media*.

Acosta, M.; Cudré-Mauroux, P.; Maleshkova, M.; Pellegrini, T.; Sack, H.; and Sure-Vetter, Y. 2019. *Semantic Systems. The Power of AI and Knowledge Graphs: 15th International Conference, SEMANTiCS 2019, Karlsruhe, Germany, September 9–12, 2019, Proceedings*. Springer Nature.

Ali, M.; Hoyt, C. T.; Domingo-Fernández, D.; and Lehmann, J. 2019. Predicting Missing Links Using PyKEEN. In *ISWC Satellites*, 245–248.

Anderson, M.; Hitlin, P.; and Atkinson, M. 2021. Wikipedia at 15: Millions of readers in scores of languages. Pew Research Center, March 2021.

Aoki, P.; Woodruff, A.; Yellapragada, B.; and Willett, W. 2017. Environmental protection and agency: Motivations, capacity, and goals in participatory sensing. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 3138–3150.

Awad, E.; Dsouza, S.; Bonnefon, J.-F.; Shariff, A.; and Rahwan, I. 2020. Crowdsourcing Moral Machines. *Communications of the ACM*, 63(3): 48–55.

Bielefeldt, A. R. 2016. First-Year Students' Conceptions of Sustainability as Revealed Through Concept Maps. In *123rd ASEE Annual Conference Proceedings, New Orleans, LA*.

Budui, R. 2014. Memory recognition and recall in user interfaces. Nielsen Norman Group, July 2014.

Cañas, A. J.; Hill, G.; Carff, R.; Suri, N.; Lott, J.; Gómez, G.; Eskridge, T. C.; Arroyo, M.; and Carvajal, R. 2004. CmapTools: A knowledge modeling and sharing environment. *Proc. of the First Int. Conference on Concept Mapping*.

Cao, Y.; Ji, X.; Lv, X.; Li, J.; Wen, Y.; and Zhang, H. 2021. Are Missing Links Predictable? An Inferential Benchmark for Knowledge Graph Completion. *arXiv preprint arXiv:2108.01387*.

Chilton, S. 2009. Crowdsourcing is radically changing the geodata landscape: case study of OpenStreetMap. In *Proceedings of the UK 24th International Cartography Conference*, 6. Citeseer.

Coeckelbergh, M. 2020. *AI ethics*. MIT Press.

Consonni, C.; Basile, S.; Manca, M.; Boratto, L.; Freitas, A.; Kovacikova, T.; Pourhashem, G.; and Cornet, Y. 2021. What's Your Value of Travel Time? Collecting Traveler-Centered Mobility Data via Crowdsourcing. *arXiv preprint arXiv:2104.05809*.

Cui, J.; and Yu, S. 2019. Fostering deeper learning in a flipped classroom: Effects of knowledge graphs versus concept maps. *British Journal of Educational Technology*, 50(5): 2308–2328.

Daley, B. J.; Durning, S. J.; and Torre, D. M. 2016. Using concept maps to create meaningful learning in medical education. *MedEdPublish*, 5.

Diakopoulos, N.; Zhang, A. X.; Elgesem, D.; and Salway, A. 2014. Identifying and analyzing moral evaluation frames in climate change blog discourse. In *Eighth International AAAI Conference on Weblogs and Social Media*.

Doan, A.; Ramakrishnan, R.; and Halevy, A. Y. 2011. Crowdsourcing Systems on the World-Wide Web. *Communications of the ACM*, 54(4): 86–96.

Dos Santos, V.; De Souza, É. F.; Felizardo, K. R.; and Vijaykumar, N. L. 2017. Analyzing the Use of Concept Maps in Computer Science: A Systematic Mapping Study. *Informatics in Education*, 16(2): 257–288.

Duhaime, E.; Olson, G. M.; and Malone, T. W. 2015. Broad Participation in Collective Problem Solving Can Influence Participants and Lead to Better Solutions: Evidence from the MIT Climate CoLab. *Collective Intelligence 2015*.

Dun, Y.; Tu, K.; Chen, C.; Hou, C.; and Yuan, X. 2021. KAN: Knowledge-aware Attention Network for Fake News Detection. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, 81–89.

Founta, A. M.; Djouvas, C.; Chatzakou, D.; Leontiadis, I.; Blackburn, J.; Stringhini, G.; Vakali, A.; Sirivianos, M.; and Kourtellis, N. 2018. Large scale crowdsourcing and characterization of twitter abusive behavior. In *Twelfth International AAAI Conference on Web and Social Media*.

Ghai, B.; Liao, Q. V.; Zhang, Y.; and Mueller, K. 2020. Measuring Social Biases of Crowd Workers using Counterfactual Queries. *arXiv preprint arXiv:2004.02028*.

Griffith, M.; Spies, N. C.; Krysiak, K.; McMichael, J. F.; Coffman, A. C.; Danos, A. M.; Ainscough, B. J.; Ramirez, C. A.; Rieke, D. T.; Kujan, L.; et al. 2017. CIViC is a community knowledgebase for expert crowdsourcing the clinical interpretation of variants in cancer. *Nature genetics*, 49(2): 170–174.

Hammon, L.; and Hippner, H. 2012. Crowdsourcing. *Business & Information systems engineering*, 4(3): 163–166.

Hu, M.; Shealy, T.; Grohs, J.; and Panneton, R. 2019. Empirical evidence that concept mapping reduces neurocognitive effort during concept generation for sustainability. *Journal of Cleaner Production*, 238: 117815.

Hung, N. Q. V.; Thang, D. C.; Weidlich, M.; and Aberer, K. 2015. Minimizing efforts in validating crowd answers. In *Proceedings of the 2015 ACM SIGMOD international conference on management of data*, 999–1014.

Jäger, G.; Zilian, L.; Hofer, C.; and et al. 2019. Crowdworking: working with or against the crowd? *J Econ Interact Coord*, 14: 761–788.

Karger, D. R.; Oh, S.; and Shah, D. 2014. Budget-optimal task allocation for reliable crowdsourcing systems. *Operations Research*, 62(1): 1–24.

Kazai, G.; Kamps, J.; and Milic-Frayling, N. 2013. An analysis of human factors and label accuracy in crowdsourcing relevance judgments. *Information retrieval*, 16(2): 138–178.

Kinchin, I. M. 2014. Concept mapping as a learning tool in higher education: A critical analysis of recent reviews. *The Journal of Continuing Higher Education*, 62(1): 39–49.

Lee, J. J.; Matamoros, E.; Kern, R.; Marks, J.; de Luna, C.; and Jordan-Cooley, W. 2013. Greenify: fostering sustainable communities via gamification. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 1497–1502.

Lemus-Rojas, M.; and Pintscher, L. 2017. Wikidata and libraries: Facilitating open knowledge. In M. Proffitt (Ed.), Leveraging Wikipedia: Connecting Communities of Knowledge (pp. 143-158). Chicago, IL: ALA Editions.

Libert, B.; Spector, J.; and Tapscott, D. 2007. *We are smarter than me: How to unleash the power of crowds in your business*. Pearson Prentice Hall.

Lim, D.; and Benson, A. R. 2020. Expertise and Dynamics within Crowdsourced Musical Knowledge Curation: A Case Study of the Genius Platform. *arXiv preprint arXiv:2006.08108*.

Maher, M. L.; and Fisher, D. H. 2012. The Role of AI in Wisdom of the Crowds for the Social Construction of Knowledge on Sustainability. In *2012 AAAI Spring Symposium Series*.

Maker, C. J.; and Zimmerman, R. H. 2020. Concept maps as assessments of expertise: Understanding of the complexity and interrelationships of concepts in science. *Journal of Advanced Academics*, 31(3): 254–297.

Margolis, J.; Estrella, R.; Goode, J.; Holme, J. J.; and Nao, K. 2017. *Stuck in the shallow end: Education, race, and computing*. MIT press.

Margolis, J.; and Fisher, A. 2002. *Unlocking the clubhouse: Women in computing*. MIT press.

Mason, W.; and Watts, D. J. 2009. Financial incentives and the performance of crowds. In *Proceedings of the ACM SIGKDD workshop on human computation*, 77–85.

Massung, E.; and Preist, C. 2013. Normification: using crowdsourced technology to affect third-party change. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems*, 1449–1454.

McMahon, C.; Johnson, I.; and Hecht, B. 2017. The substantial interdependence of Wikipedia and Google: A case study on the relationship between peer production communities and information technologies. In *Eleventh International AAAI Conference on Web and Social Media*.

MediaWiki. 2021. Wikidata query Service/User manual. Https://www.mediawiki.org/wiki/Wikidata_Query_Service/User_Manual, Accessed On 8/1/2021.

Mitra, T.; Hutto, C. J.; and Gilbert, E. 2015. Comparing person-and process-centric strategies for obtaining quality data on amazon mechanical turk. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 1345–1354.

Nielsen, F. Å. 2019. Ordia: A Web Application for Wikidata Lexemes. In *European Semantic Web Conference*, 141–146. Springer.

Nielsen, F. Å.; Mietchen, D.; and Willighagen, E. 2017. Scholia, scientometrics and Wikidata. In *European Semantic Web Conference*, 237–259. Springer.

NOAA. 2020. Citizen science and crowdsourcing. https://www.noaa.gov/office-education/citizen-science-crowdsourcing. Accessed: 2020-11-10.

Novak, J. D. 1990. Concept mapping: A useful tool for science education. *Journal of research in science teaching*, 27(10): 937–949.

Novak, J. D.; Gowin, D. B.; and Kahle, J. B. 1984. *Learning How to Learn*. Cambridge University Press.

Noy, N.; Gao, Y.; Jain, A.; Narayanan, A.; Patterson, A.; and Taylor, J. 2019. Industry-scale knowledge graphs: lessons and challenges. *Communications of the ACM*, 62(8): 36–43.

Pellissier Tanon, T.; Vrandečić, D.; Schaffert, S.; Steiner, T.; and Pintscher, L. 2016. From freebase to wikidata: The great migration. In *Proceedings of the 25th international conference on world wide web*, 1419–1428.

Piccolo, L.; Fernández, M.; Alani, H.; Scharl, A.; Föls, M.; and Herring, D. 2016. Climate change engagement: Results of a multi-task game with a purpose. In *Tenth International AAAI Conference on Web and Social Media*.

Plotnick, E. 1997. *Concept mapping: A graphical system for understanding the relationship between concepts*. ERIC Clearinghouse on Information and Technology Syracuse, NY.

Qi, G.; Gao, H.; and Wu, T. 2017. The research advances of knowledge graph. *technology intelligence engineering*, 3(1): 4–25.

Salehi, N.; McCabe, A.; Valentine, M.; and Bernstein, M. 2017. Huddler: Convening Stable and Familiar Crowd Teams Despite Unpredictable Availability. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, CSCW '17, 1700–1713. ISBN 9781450343350.

Satyapanich, T.; Ferraro, F.; and Finin, T. 2020. Casie: Extracting cybersecurity event information from text. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 8749–8757.

Schwartz, B. 2014. Google's Freebase to close after migrating to Wikidata: Knowledge graph impact? *Search Engine Roundtable*, 17.

Segalàs, J.; Ferrer-Balas, D.; and Mulder, K. F. 2008. Conceptual maps: measuring learning processes of engineering students concerning sustainable development. *European Journal of Engineering Education*, 33(3): 297–306.

Segalàs, J.; Ferrer-Balas, D.; and Mulder, K. F. 2010. What do engineering students learn in sustainability courses? The effect of the pedagogical approach. *Journal of Cleaner Production*, 18(3): 275–284.

Shallcross, D. C. 2016. Concept maps for evaluating learning of sustainable development. *Journal of Education for Sustainable Development*, 10(1): 160–177.

Shaw, A. D.; Horton, J. J.; and Chen, D. L. 2011. Designing incentives for inexpert human raters. In *Proceedings of the ACM 2011 conference on Computer supported cooperative work*, 275–284.

Shenoy, K.; Ilievski, F.; Garijo, D.; Schwabe, D.; and Szekely, P. 2021. A Study of the Quality of Wikidata. *Journal of Web Semantics*, 100679.

Siew, C. S. 2019. Using network science to analyze concept maps of psychology undergraduates. *Applied Cognitive Psychology*, 33(4): 662–668.

Silberman, M. S.; Tomlinson, B.; LaPlante, R.; Ross, J.; Irani, L.; and Zaldivar, A. 2018. Responsible research with crowds. *Communications of the ACM*, 61(3): 39–41.

Spitz, A.; Dixit, V.; Richter, L.; Gertz, M.; and Geiß, J. 2016. State of the union: A data consumer's perspective on wikidata and its properties for the classification and resolution of entities. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 10.

Steiner, H. H. 2016. The Strategy Project: Promoting Self-Regulated Learning through an Authentic Assignment. *International Journal of Teaching and Learning in Higher Education*, 28(2): 271–282.

Tan, S.; Erdimez, O.; and Zimmerman, R. 2017. Concept Mapping as a Tool to Develop and Measure Students' Understanding in Science. *Acta Didactica Napocensia*, 10(2): 109–122.

Turki, H.; Jemielniak, D.; Taieb, M. H.; Gayo, J. L.; Aouicha, M. B.; Banat, M. a.; Shafee, T.; Prud'Hommeaux, E.; Lubiana, T.; Das, D.; et al. 2020. Using logical constraints to validate information in collaborative knowledge graphs: a study of COVID-19 on Wikidata. *Semantic Web Journal*.

Vrandečić, D. 2013. The rise of Wikidata. *IEEE Intelligent Systems*, 28(4): 90–95.

Vrandečić, D.; and Krötzsch, M. 2014. Wikidata: a free collaborative knowledgebase. *Communications of the ACM*, 57(10): 78–85.

Waagmeester, A.; Stupp, G.; Burgstaller-Muehlbacher, S.; Good, B. M.; Griffith, M.; Griffith, O. L.; Hanspers, K.; Hermjakob, H.; Hudson, T. S.; Hybiske, K.; et al. 2020. Science Forum: Wikidata as a knowledge graph for the life sciences. *Elife*, 9: e52614.

Walshe, N. 2008. Understanding students' conceptions of sustainability. *Environmental Education Research*, 14(5): 537–558.

Wikidata. 2021a. Wikidata Statements. Https://en.wikipedia.org/wiki/WikidataStatements, Accessed On 8/1/2021.

Wikidata. 2021b. Wikidata Stats. Https://wikidata-todo.toolforge.org/stats.php Accessed On 8/1/2021.

Wikimedia. 2021. Wikistats - statistics For Wikimedia projects. Https://stats.wikimedia.org/#/wikidata.org, Accessed On 8/1/2021.

WorldBank. 2021. Higher Education. World Bank, https://www.worldbank.org/en/topic/tertiaryeducation1, Accessed On 1/13/2021.