RESEARCH ARTICLE

LIS Journals' Lack of Participation in Wikidata Item Creation

Eric Willey
Illinois State University

Susan Radovsky
Harvard Library

There are many items in Wikidata representing scholarly articles. However, these items have been created mostly by volunteer Wikidata editors and not systematically by journal publishers or editors, which can lead to gaps and inconsistencies in the datasets. This article presents findings from a survey investigating practices of library and information studies (LIS) journals in Wikidata item creation. Believing that a significant number of LIS journal editors would be aware of Wikidata and some would be creating Wikidata items for their publications, the authors sent a survey asking 138 English-language LIS journal editors if they created Wikidata items for materials published in their journal and follow-up questions. With a response rate of 41 percent, respondents overwhelmingly indicated that they did not create Wikidata items for materials published in their journal and were completely unaware of or only somewhat familiar with Wikidata. Respondents indicated that more familiarity with Wikidata and its benefits for scholarly journals as well as institutional support for the creation of Wikidata items could lead to greater participation; however, a campaign of education about Wikidata, documentation of benefits, and support for creation would be a necessary first step. The article presents and discusses the results of the survey, but the conclusions that can be drawn are minimal; therefore, the authors also discuss the benefits of creating Wikidata items for LIS journals as a first step in this educational campaign for editors and publishers.

Introduction

A sister project to Wikipedia, Wikidata was created with the goal of providing a machine- and human-readable database for linked open data objects (i.e., Wikidata items), which can describe anything that meets Wikidata's (much broader than Wikipedia) notability guidelines. Now over ten years old, Wikidata is the largest public knowledge graph in the world, with over one hundred million items created by over half a million editors (Vrandečić, Pintscher, and Krötzsch 2023). The project has seen considerable growth since inception and is often used in other Wikimedia and third-party projects, such as knowledge graphs in Wikipedia articles and search results.[1] It has even absorbed other knowledge graphs such as Freebase, which Google cancelled in 2016, migrating the information in it to Wikidata (Vrandečić, Pintscher, and Krötzsch 2023).

As the central storage repository for the structured data from all Wikimedia projects, Wikidata has increasingly become a way for galleries, libraries, archives, and museums (GLAMs) to share data about their collections in a linked open data environment (Lemus-Rojas et al. 2022; Kent 2019; Yon and Willey 2021). As discussed by Allison-Cassin and Scott (2018), Wikidata has "provid[ed] a ready-made platform for any person or organization that wants to create, publish, and use LOD, including libraries." The benefits of Wikidata as a global knowledge base include:

---

[1] For a general overview, see *Wikidata for Scholarly Communication Librarianship* by Jere Odell, Mairelys Lemus-Rojas, and Lucille Brys (2022).

- Interoperability. Wikidata is designed to serve as a database for various applications and platforms, such as openArtBrowser, Cita, Entity Explosion, Scholia, and others ("Wikidata:Tools/Visualize data" 2023).
- Disambiguation. By assigning unique identifiers to each item and property, Wikidata differentiates between entries described with the same or similar terms. These unique identifiers also make the metadata machine readable, which allows AI, bots, scripts, and other software to provide more thorough and accurate answers to questions, provide more context, link to related resources, and potentially prioritize more accurate information in search results (Kent 2019).
- Data enrichment. Through structured properties and relationships, Wikidata allows for links and context to be created between different items in the knowledge base. WDQS, the Wikidata Query Service, then allows users to pull data and explore relationships between all these linked items.
- Adherence to linked open data principles. Wikidata promotes the linking and sharing of metadata across the web. The effects are similar to those provided by making the metadata machine readable. According to Zuiderwijk, Jeffery, and Janssen, linked open data, with the appropriate architecture, can yield benefits such as "improving findability, accessibility, storing, preservation, analysing, comparing, reproducing, finding inconsistencies, correct interpretation, visualizing, linking data, assessing and ranking the quality of data and avoiding unnecessary duplication of data" (2012, 222). But the open nature of the data goes even further: making metadata freely available for use and reuse allows editors and users to build upon each other's work rather than duplicate effort behind paywalled silos or proprietary formats that not all users have access to.
- Sustainability. Wikidata is volunteer driven and does not rely on proprietary software. The data is also free and open, meaning that it can be linked to, reused, and distributed by anyone without a fee.
- Multilingualism. Wikidata supports translation between languages by design. While there are over three hundred versions of Wikipedia in different languages, there is only one version of Wikidata from which all Wikipedias pull data, enabling editors around the world to collaborate and ultimately providing more equitable access to knowledge.

For GLAMs, Wikidata affords the opportunity to make their collections more visible, accessible, and useful to a global community of users and to establish connections between collections held at different institutions.

Similarly, Wikidata offers multiple benefits for scholarly journals. The knowledge base contains a large amount of existing bibliographic data that can be used to connect journals to the articles they publish and to situate those articles within a larger scholarly context. Adding Wikidata items for scholarly articles improves visibility and discoverability of journal content, especially by providing additional access points to articles beyond author, title, and subject or keywords. This data allows users to explore relationships among articles and among articles and other entities such as authors, institutions, awards, sponsors, and publishers, etc. Since many of these entities exist in Wikidata already, properties and their corresponding items for scholarly articles can often be added without needing to be created first. Journals thereby benefit from the work of the larger Wikidata community. Through this web of linked data, articles in GLAM journals could also be linked (directly or indirectly) to Wikidata items containing collection data about their topics, and vice versa.

Wikidata also contains properties for various linked data persistent identifiers (e.g., DOIs, Directory of Open Access Journals IDs, ORCID iDs, etc.) and for citations (through the *cites work* and *cited by* properties), aligning it with other open metadata initiatives working to make bibliographic data open and reusable, such as the Initiative for Open Citations (I4OC), OpenCitations Corpus, and OpenAlex. The citation and bibliographic metadata in Wikidata is also integrated into the larger Wikimedia structure under the WikiCite project, which includes creating an open bibliographic database in Wikidata among its goals ("Wikidata:WikiCite" 2023).[2] Because Wikipedia relies on verifiability for its articles, having a reliable, easy-to-use source for citations is tremendously helpful to editors and the project. Current goals for WikiCite include importing data for academic works, resolving author name strings (text that shows an author's name but is not linked data), disambiguating author names, curating sets of items via Wikidata lists, and adding subjects to all scientific articles ("Wikidata:WikiCite" 2023). This has increased the amount of metadata describing scholarly articles by a huge amount.

---

[2] The relationship between Wikidata and WikiCite is of interest to authors and publishers of scholarly materials, which measure impact and success partly on the number of times they are cited by other works. The WikiCite Initiative was created with the intention of creating a "universal repository of open bibliographic data of citable sources within Wikidata" ("WikiCite/Shared Citations" 2023). The structured, machine-readable format of Wikidata makes it ideal for automated citation creation (and it is inherently multilingual, allowing for greater diversity and international representation of scholarship), which can contribute greatly to raising the visibility and citation rates of scholarly articles. A 2016 report on WikiCite offers further detail on Wikidata's critical role in the project (Taraborelli et al. 2016).

However, although many Wikidata items representing scholarly articles exist, they are rarely created by the journal publishers or editors in a systematic way.[3] Items for scholarly articles are mostly created by volunteer Wikidata editors, often as a part of the WikiCite initiative (as noted above), or by bots (Lemus-Rojas et al. 2022), which can lead to gaps and inconsistencies in the datasets. Not all articles from a journal's entire publication history always have Wikidata items, and some Wikidata items for articles have metadata that others do not. Ideally, creating items for scholarly articles would include creating or including existing Wikidata items for associated entities, such as article authors, as well. Yet, although sometimes Wikidata items for authors are created and used, other times the authors' names are listed as text strings that are not linked data. Having a complete corpus with consistent data would improve the size and completeness of the dataset for researchers and allow for more thorough indexing of journal articles by search engines.

Recognizing the benefits of Wikidata both for GLAMs and for scholarly journals, the authors of this article, who work in the field of library and information science and are LD4 Wikidata Affinity Group coordinators, surveyed LIS journal editors to establish how many LIS journals are currently creating Wikidata items for the articles they publish. We assumed that, as part of the GLAM sector, respondents would be at least generally familiar with Wikidata, especially since the library community has embraced other open metadata initiatives such as open citation and persistent identifier projects. However, when surveyed, no LIS journal editors answered that they or their publisher were systematically creating Wikidata items for their materials, and many of them were not familiar with Wikidata at all.

This article reports the findings from our survey and suggests that a potential way to standardize Wikidata item creation for journal articles would be for editors or publishers to create Wikidata items for all articles at the time of publication (and for previously published articles as well). As Kemp, Dean, and Chodacki (2018) point out, improving the quality of metadata for their publications is a priority for publishers,[4] and we contend that investing in the creation of linked data in Wikidata would serve this goal. Institutional support can also often provide consistency in workflows and a degree of succession planning (helping to ensure the long-term sustainability of the systematic approach) that volunteer work sometimes does not, and volunteers could still enhance and improve Wikidata items beyond what is provided by publishers or editors. Our survey suggests that demonstrating the value of adding metadata in Wikidata could increase engagement in creating Wikidata items, and we hope that this article will encourage journal publishers and/or editors to consider systematic Wikidata item creation for their articles, ultimately offering a more complete and consistent search experience for users and improving the world of linked data.

Literature Review

We found only one article about an LIS journal editor or publisher creating Wikidata items, so this literature review includes scholarship that discusses the existing dataset of scholarly articles in Wikidata more broadly as well as methods for creating Wikidata items for scholarly articles. The literature review began by searching Google Scholar for the keywords *LIS* (and variations) and *Wikidata scholarly articles.* Searches for *publishers Wikidata* as well as *scholarly journals Wikidata* were also conducted. As few relevant articles were being discovered, searches for *publishers metadata* and *scholarly communications* were also made. When an article of interest was found, the authors checked Google Scholar for works that cited the article and might be relevant and the bibliography of the article itself. The authors also happened to know of some articles through their personal interests and work. The authors found very few relevant articles describing projects by LIS journals to create Wikidata items for their materials, but (optimistically and erroneously) believed that at least some journals were creating Wikidata items but had not published articles on the process; therefore, the initial literature review focused on that theme. Additional articles were later incorporated based on peer reviewer and editorial feedback.

---

[3] The Wikidata Terms of Use/FAQ on paid contributions without disclosure ("Policy:Terms of Use" 2023) allows for the creation and editing of items describing articles by their editors or publishers. Although paid editing must be disclosed by the editor, existing projects by academic librarians to add items for theses and dissertations by members of their institution, such as the exemplary work of Adam Schiff, Crystal Yragui, and their colleagues at the University of Washington ("Wikidata:WikiProject PCC Wikidata Pilot/University of Washington" 2023), are known and accepted. Items describing journal articles are also added regularly and do not raise issues of notability.

[4] In their literature review of metadata in scholarly communications, Gregg et al. (2019) conclude that publishers are increasingly recognizing the return on investment that metadata provides and are likely to invest in it more heavily. This review is focused on descriptive metadata but may be applicable to linked data as well. Gregg et al. (2019) also cite the interesting idea from Emily Alinder Flynn (2013) that vendors' products would be more valuable to libraries if they included higher quality metadata. Whether vendors would be inclined to create this metadata in Wikidata and release it as open access is uncertain, but the large amount of existing data in Wikidata and support from the community might make it more viable than building proprietary systems.

Scholarly Articles in Wikidata

As far as the authors know, there is only one example of a published, peer-reviewed article describing an LIS journal taking direct responsibility for systematically creating Wikidata items for all their published articles: the Italian journal *JLIS.it*, which used an automated process to create Wikidata items for its scholarly articles from the journal's inception in 2010 through 2019 (Bianchini 2021).[5] Generally, however, the number of Wikidata items for scholarly articles has increased since Nielsen, Mietchen, and Willighagen presented their findings on the completeness of Wikidata in relation to scholarly bibliographic data in 2017. At that point, they found that "journals and universities are well represented" in Wikidata but "far less covered are individual articles, individual researchers, university departments and citations between scientific articles" (Nielsen, Mietchen, and Willighagen 2017, 241). Observing that "most of the scientific articles in Wikidata are claimed to be an *instance of* (P31) the Wikidata item *scientific article* (Q13442814)," they identified 2,380,009 instances of *scientific article* in the knowledge base (Nielsen, Mietchen, and Willighagen 2017, 241).[6] More recently, a [query] of Wikidata items for *scholarly article* (and subclasses) by Cobb (2020) in October 2020 showed 38.7 million items. Repeating this search on October 29, 2023, showed a tally of 43,898,288 items.

In terms of who is creating items for scholarly articles, Lemus-Rojas et al. explain that the Wikidata "community has taken an interest in increasing the representation of bibliographic data" in the knowledge base, "with a focus on scholarly articles," and that "many of these contributions are the efforts of WikiCite, an initiative and a community that aims at building an open citation database in Wikidata" (2022, 3). However, systematic efforts at creating Wikidata items for scholarly articles tend to be institution focused rather than journal focused. For example, Lemus-Rojas and Odell (2018) describe a pilot project to create Wikidata items for the faculty of the Indiana University Lilly Family School of Philanthropy, their publications, and co-authors of the publications. Scholia (a website that generates profiles for people, journals, and organizations based on information from Wikidata) is used to display the faculty profiles they created.[7] Upon completion of the pilot, they examined ways to add data from other systems their library already had and facilitated Wikidata edit-a-thons to increase awareness and knowledge of Wikidata in the Indiana University–Purdue University Indianapolis (IUPUI) library. The researchers identified two challenges to creating Wikidata items: the need to generate institutional buy-in and the need to find a source of labour both to do manual entry where necessary and to prepare data for batch processing into Wikidata items (Lemus-Rojas and Odell 2018). Building on that research, Lemus-Rojas et al. (2022) focused next on populating Wikidata with data about women scholars employed by IUPUI, arguing that building profiles of scholars through Wikidata can address information inequities as well as support open knowledge goals.

Overall, this scholarship suggests a positive trend in the addition of items for scholarly articles in Wikidata. However, it also suggests that the impetus for adding these items is often to improve the representation of specific scholars, groups of scholars, or institutions—rather than journals—in Wikidata and that journals may be better represented in the knowledge base if they undertook the creation of Wikidata items for their content themselves.

Methods of Adding Wikidata Items for Scholarly Articles

There are several possible methods for adding items for scholarly articles to Wikidata. As Lemus-Rojas et al. explain, "most of the article contributions are being made through bots—tools used for making automated contributions without the need for human intervention," but users can also add article data "either by using external tools or by manual editing" (2022, 3). Alves, Burley, and Peschanski (2021) describe a largely automated process for adding article metadata for the journal *Anais do Museu Paulista* to Wikidata. This process relies on importing article metadata into Zotero via unique identifiers and then exporting QuickStatements, which can add information to Wikidata automatically instead of going field by field manually. They also describe a then-recent software addition to Wikipedia that allows editors to create a cross reference through the QID for the referenced article's Wikidata item. This helps improve consistency and automates the citation process (and in the future may even allow automatic updates of citations in Wikipedia when a Wikidata item is updated) but can also then be reused for citations in Wikipedia articles in different languages (Alves, Burley, and Peschanski 2021). Based in the Global South, their project demonstrates how low-barrier tools and processes can help increase diversity in scholarly content in Wikidata (Alves, Burley, and Peschanski 2021).

---

[5] The body of the article is in Italian, which neither of the current authors read. All information about the article and work is taken from the English language abstract, with apologies for any inaccuracies or omissions.

[6] The primary label for such an article at the time was *scientific article*, but it has since changed to *scholarly article*, with *scientific article* included under "Also known as."

[7] Bianchini (2021) notes that, since librarians at the University of Salerno have also added metadata to Wikidata for articles in the journal *Bibliothecae.it*, having Wikidata items for *JLIS.it* allows for quantitative analysis of Italian LIS literature in Scholia.

For their work on increasing the representation of women editors of periodicals in Wikidata, Thornton et al. (2023) discuss a different process in which they exported metadata from a database into CSV files, checked with a Python script, and then used OpenRefine (with a reconciliation extension specifically for Wikidata) to reconcile people, organizations, and periodicals in Wikidata (Thornton et al. 2023). From there they used the WDI, or WikidataIntegrator (a collection of Python scripts) to write their reconciled data to Wikidata and the django-wikidata-api to minimize the risk of creating duplicate items (Thornton et al. 2023).

Given the wide variety of possible approaches to Wikidata item creation, the authors had hoped to discover articles describing projects by journal editors or publishers to systematically create Wikidata items for materials they published, their reasons for doing so, and any benefits or outcomes of that work through the literature review. When this information was largely not found, the authors were not certain if this work was not being done or being done but not discussed in the literature or Wikidata community. Therefore, they developed a survey with two potential audiences in mind: 1) publishers/editors who were systematically creating Wikidata items but not publicizing their methods or the results of their work and 2) publishers/editors who were not creating Wikidata items.

Methods

Survey Design

The intention of our fourteen-question survey ([Appendix 1](#)) was to document how and why some LIS journals were creating Wikidata items, why others were not, and then use the responses from each group to generate a convincing case and way forward for those who were not. We saw involvement with the LD4 Wikidata Affinity Group from members of other GLAM institutions, and we thought that by focusing on LIS journals at least some respondents would be familiar with Wikidata and even be adding items for materials published in their journal. This was not the case, and questions four through eight (which focused on why and how respondents created Wikidata items) did not apply to any respondents. The answers to these questions would have been used to generate reasons to create Wikidata items and recommendations for how to do so for respondents that did not create Wikidata items.

Responses to questions twelve through fourteen were similarly of low use in determining how publishers or editors worked with Wikidata. Question twelve asked if respondents would like to increase their involvement in creating Wikidata items, but the authors had overestimated respondents' familiarity with the platform. As a large majority of respondents were not at all or only somewhat familiar, there were only four responses to this question. Question thirteen, regarding if the journal was currently publishing new material, was included in case the COVID-19 pandemic had and continued to disrupt publishing for a significant numbers of journals, but this was not the case. Question fourteen was about the open access status of the journal, with the intention of seeing if there was a correlation between open access publishing and adoption of Wikidata. It is conceivable that open access journals might also benefit the most from creating Wikidata items because they can link directly to content rather than paywalled materials.

Distribution

Authors used the "Peer Reviewed LIS Journals" list created by Selinda Berg and updated by Kristin Hoffmann (2022) at Western University to identify potential survey participants and located email contact information for LIS journal editors by searching each journal's web page. This list was chosen because it was convenient, did not restrict itself to high-impact or "top" LIS journals, and had information about who provided updates for it in case there were any questions. The list was of English-language journals, which resulted in *JLIS.it* (discussed in the literature review) not being included. Nearly all journals listed more than one email for their various staff and contributors, and emails for people described as *contact person* or *editor* (or the closest equivalent) were chosen. Only one person per journal was directly emailed the link to the Qualtrics survey, and the email specified which journal's practices they were being asked to describe because some people edit multiple journals which might have different practices. The email requested that it be forwarded if necessary, but asked that only one reply per journal be made. Replies were anonymized and not associated with specific journals or recipients of the emails. The data was collected anonymously to comply with the General Data Protection Regulation (GDPR).

There were 142 journals on the initial list, but four were not contacted because they had not recently published articles and did not have contact information. These journals were *Behavioral & Social Sciences Librarian* (had not published since volume 36, 2017), *Community & Junior College Libraries* (had not

published since volume 23, 2017), *Library & Information History* (had not published since volume 35, 2019), and *Reference and User Services Quarterly*.[8]

Results

Of the 138 journals contacted, fifty-seven replied, giving a response rate of 41 percent. Question one was a standard consent form. For question two, all fifty-seven survey participants indicated their level of familiarity with Wikidata. Twenty-eight of the fifty-seven respondents (49 percent)[9] had no familiarity with Wikidata, while twenty-five respondents (44 percent) were somewhat familiar with Wikidata, leaving only four respondents (7 percent) who were very familiar with Wikidata. This means that 93 percent of those surveyed had only a moderate exposure, if any, to Wikidata, and this certainly informs the rest of our survey results to a large degree.

All fifty-seven of our participants responded to question three, concerning whether the respondent or anyone else on their staff adds metadata about materials published in their journal to Wikidata. In this case, fifty-two of the fifty-seven participants (91 percent) responded in the negative. The remaining five participants (9 percent) noted that they were aware of others outside the journal and publisher creating items for their materials, though no other details were provided in that section of the survey.

Questions four through eight were only displayed if respondents answered *yes* to question three, "Do you or a member of your journal staff add metadata about materials published in your journal to Wikidata?" Because no respondents replied that they were creating or editing Wikidata items, there were no answers to these questions.

Question nine, "Why do you not add information about your journal's material to Wikidata?" (respondents could check all reasons that applied), received ninety-four total responses. The most common response ($n = 39$) was that respondents did not know it was an option. *Lack of time or other resources* ($n = 23$) and *don't know how to add to Wikidata* ($n = 15$) were also common responses. Although not selected by a large number of respondents, *Wikidata can be edited by anyone/lack of control over the metadata* ($n = 4$), *no benefit to adding items to Wikidata* ($n = 6$), *concerns over author privacy or potential misuses of metadata* ($n = 2$), and *other* ($n = 5$) were also selected. The free-text responses to *other* indicated a lack of awareness of Wikidata, adding Wikidata items being outside the scope of the editor's duties, not being clear on the benefits of adding Wikidata items, and not receiving any datasets that could be put into Wikidata (which may indicate that the respondent is not entirely aware of what sort of information is added to Wikidata). The first and third most common responses were related to a lack of knowledge about Wikidata and how to add items, and the second most common response was lack of time or other resources. The fact that there were only six responses that indicated the respondent felt there was no benefit to adding items to Wikidata may indicate that there is a general feeling of goodwill towards the platform or at least the belief that metadata has inherent value. However, all the respondents were editors of journals somehow associated with librarianship, and this attitude may not generalize to editors of journals related to other disciplines (and, of course, six responses from editors of LIS journals did feel there was no benefit in adding items to Wikidata).

Respondents were asked to select all options that applied to question ten ("What resources would help you add information about your materials to Wikidata if you did wish to do so?"), and the authors received 131 responses. The most common responses were *training* ($n = 27$), *user studies on the benefit of adding information about materials* ($n = 25$), *documentation on best practices* ($n = 24$), *recommended workflows* ($n = 24$), and *customized workflows for my journal's materials* ($n = 19$). A handful of respondents also selected *productivity tools* ($n = 5$) and *other* ($n = 7$).

Based on the response to question three, to which only five participants noted that they were aware of people not affiliated with the journal creating metadata for materials published in the journal to Wikidata, it is not surprising that only four participants responded to question eleven, which asked, "Who is creating Wikidata items for materials in your journal (select all that apply)?" Three of the four respondents (75 percent) indicated *that individual volunteer(s) unaffiliated with the journal or publisher* were creating related metadata. The remaining single respondent chose *other*, but did not elaborate.

Question twelve, "Would you like to increase your direct involvement in creating Wikidata items for materials in your journal?," also had a low response rate. But although only four respondents answered, the results point in a hopeful direction: two respondents replied with a *yes*, and two with a *maybe*. Zero respondents

---

[8] The journal's "About" page states, "*RUSQ* is not accepting submissions at this time. The RUSA Publications Taskforce will be making recommendations regarding all RUSA publications in the spring of 2022 and we will share the results of that report here at that time. Thank you for your interest in *RUSQ* (9/19/2021)."
[9] All percentages have been rounded.

said *no*, they would not be interested in increasing their direct involvement in creating Wikidata items for their materials.

Respondents were also asked if the journal they edited was currently publishing new material. Out of fifty-three total responses, only three indicated they were not publishing new material at the time of the survey. The remaining fifty indicated they were currently publishing new materials. This question was asked primarily in case journals were paused or on unofficial hiatus due to the COVID pandemic or other reasons (which would indicate why they were not creating Wikidata items for new materials at least), but a clear majority of journals were publishing new material at the time of the survey. It is also possible that editors of journals that were not publishing new materials did not see the email for the survey or chose not to respond.

Finally, respondents were asked, "Is this journal Open Access (materials can be viewed without a paid subscription or other financial charge)?" Of fifty-three total responses, twenty-six (49 percent) indicated all materials published by the journal they edited could be accessed without payment. A further twenty-five respondents (47 percent) indicated some portion of their materials were open access or authors were allowed to make pre-print copies accessible, and only two respondents (4 percent) indicated that all of their materials were paywalled.

Discussion

Survey results clearly indicate that educating LIS journal editors and publishers about Wikidata and the benefits it can offer would be useful. Fortunately, there is already specific documentation, such as "Wikidata:WikiProject Periodicals" (2022), although promoting it as part of education efforts might be beneficial. Workflows for specific journals, or at least a generic template of pick-and-choose options, might also lower the barriers to entry for editors to begin adding Wikidata items for the journal they edit (or convince their publishers or platforms such as Open Journal Systems to provide support to do so). A low-barrier-to-entry option would be especially useful in alleviating the concerns about where staff time would come from raised by survey respondents and documented by Lemus-Rojas and Odell (2018). Automation techniques such as those discussed by Alves, Burley, and Peschanski (2021) would help reduce the amount of staff time required, although they would require training.

It might also be possible to convince authors to do some of the work in creating their own items, further reducing workload for editors. Wilkins et al. (2022) discuss the willingness of authors to enter information about their study methods and results as structured data (although not as Wikidata items) during the article submission process. The researchers found that directly asking authors to provide this information about their article through pre-generated templates might be viable. If authors were willing to provide information about themselves and/or their articles to include in Wikidata items, that would lower the amount of labour required from editors, other journal staff, or publishers. This information could include institutional affiliations, professional memberships, social media accounts, or nearly anything else desired. If authors desired, it could also include sensitive information such as race or ethnicity, sexuality, or other traits for living persons that present complicated ethical issues when included as linked data without an author's consent. It would likely also be more efficient than (for example) sending all the authors of an article a questionnaire, asking them to fill it out, then having a journal editor or other worker translate that questionnaire into an item on Wikidata.

However, no matter how low the barriers, survey results show that demonstrating a clear benefit to creating Wikidata items will be necessary for buy-in from editors and publishers. An interview or discussion with Bianchini (2021) about any benefits he has been able to identify from this project would likely be valuable. Zuiderwijk, Jeffery, and Janssen's (2012) research could also provide a convenient checklist for how creating Wikidata items can benefit journals. An evaluation and demonstration of those benefits would be needed, but this research can provide at least an initial direction. Editors and publishers may also be motivated if they become aware that some of their articles are already being described in Wikidata through projects like that of Lemus-Rojas et al. (2022). Because projects like these focus on existing faculty or sub-groups of existing faculty, they naturally add items for more recently published articles. Journals with rich historical catalogues of materials may wish to add items for back issues of journals to promote those materials as well. For example, the *British Journal of Educational Technology* began publishing in 1970 but only has Wikidata items for two articles published before 1983 and a considerable number of items for articles published after 2002 (Wikidata Query Service 2023).

Overall, survey results show that editors and publishers have not systematically added Wikidata items for their materials because no case has been made showing why and how they should do so. While the volunteer and hobbyist Wikidata community has made enormous contributions to the linked data ecosystem, institutional support may remain largely absent until the impact of creating Wikidata items is documented and clear procedures are provided, at least as a beginning step for journals to calculate the time

commitment of Wikidata item creation. Since survey responses identified time and resources as barriers, it also seems logical that a strong use case will be necessary if the amount of work involved in the suggested procedure is high, but a process involving less work (i.e., a largely automated process) might require fewer demonstrated benefits to seem like a reasonable investment of time and money.

Conclusions and Recommendations

Given the widely varied needs, resources, and practices of the journal publishing landscape, a one-size-fits-all approach will not be effective in encouraging editors or publishers to create Wikidata items for their materials. What can be done is a one-size-fits-some or perhaps a one-size-fits-most approach that will cover a significant portion of journals and accommodate different approaches for the remainder. Therefore, several recommendations follow with the intention that they can be applied where wanted or needed, and other options will be available if they are not. These recommendations are also based on our survey of LIS journals and may or may not generalize to journals in other disciplines.

First of all, lack of knowledge about Wikidata is a major obstacle. Establishing an educational program for publishers and editors of journals that explains what Wikidata is and what it can do will be a necessary first step if they are going to allocate resources to creating Wikidata items. This work would likely fall to members of the Wikidata community and might be done by presenting at conferences, publishing journal articles, or promoting public awareness through campaigns like Wikipedia edit-a-thons. Much of the information on how to create items that would be required in these presentations has already been worked out by the Wikidata community, such as the previously mentioned "Wikidata:WikiProject Periodicals" (2022) group. As awareness grows, some editors and publishers may see intrinsic value in this work, but in any case education will be a necessary first step before adoption.

Education efforts cannot focus exclusively on how to create Wikidata items but must also communicate the benefits of Wikidata that differentiate it from other systems or platforms. As noted in the introduction, the library community has widely embraced open metadata initiatives such as open citation and persistent identifier projects. It would be advisable to stress what Wikidata can provide beyond adding another citation to the web: rich knowledge graphs documenting relationships between articles, creators, and other entities; data which is both machine and human readable, allowing search engines and people to read it; and a rich volunteer community exploring new ways to reuse the data such as Listeria, a program which can automatically generate lists for Wikipedia about articles or authors (McAndrew and Strathmann 2021). Because Wikidata is open access, it can be promoted as a way for a journal to promote its articles, enhance visibility for authors and research, and supplement traditional indexing approaches with no direct work on the journal's part beyond the initial item creation. Editors and publishers face the same financial and time constraints as many in academia do, so it is necessary to persuade them that a low initial investment of time and coordination with the volunteer community can yield benefits for their articles and authors.

Existing avenues to raise initial awareness of Wikidata include sessions, presentations, and workshops at appropriate conferences. Publishing more scholarship on Wikidata will also raise awareness among journal editors and peer reviewers as well as (depending on the findings of the research, hopefully) demonstrate its value and impact. A robust network of support for creating Wikidata items exists, but it may be beneficial to explore a similar network to support scholarship about Wikidata. Applications of Wikidata, such as the website Scholia, can be used to demonstrate possible uses for the data. The LD4 Wikidata Affinity Group (2023) has supported the Wikidata community since 2019 with group calls featuring volunteer speakers on various topics, since 2020 with working hours which allow people to create Wikidata items while asking other participants questions and discussing processes, and since 2021 for those interested in Wikibase and WBStack. These presentations are open to all and are often recorded, with the recordings made openly available afterward. Coordination of the group is also open to all, with volunteers free to decide how much they would like to be involved. Anecdotally, participants are usually in GLAM institutions, and the ongoing sessions and recordings could be promoted as an existing educational resource for editors or publishers seeking more information generally or on specific topics.

A large portion of the editors surveyed indicated that their journal published materials as open access to one extent or another. This suggests that stressing the open access nature of Wikidata might appeal to them or align well with their institutional mission. This could also be contrasted with citation data generated by Google and Google's history of cancelling support for some of their services such as Google Reader, Google Stadia, and Google Labs ("Category:Discontinued Google services" n.d.). Canned searches demonstrating that users can easily identify articles or journals which are open access could also be used to demonstrate value, especially for hybrid open access journals who might wish to sort restricted and publicly available material.

Once the educational foundation is in place, it will be helpful if not necessary to develop best practices for publishers and editors to refer to when adding Wikidata items. One example might be for creating metadata using automated processes, such as importing information to Zotero via DOIs or the Zotero Connector plugin and exporting Quickstatements. This sort of partially or fully automated process (depending on the type of metadata desired) is something that can be developed by the Wikidata community and third parties and then implemented by publishers to ensure comprehensive and prompt creation of items for their journals. Bots can also be developed and run on Wikidata, and "can add interwiki links, labels, descriptions, statements, sources, and can even create items, among other things" ("Wikidata:Bots" 2023).

Whether the process developed by Thornton et al. (2023) would be a useful approach for journals depends on the size and scope of the project. Adding Wikidata items for ten thousand articles from the entire history of a journal is a wildly different project than just adding five articles for a new issue. Generally, it seems likely that an initial process that is automated as much as possible to get existing articles into Wikidata, followed by a less automated process for adding new materials in small batches, would be a balanced approach. The Wikidata community itself uses a wide variety of approaches to add items for journal articles. Some try to automate the process as much as possible to handle large bulk uploads, while editors who are only creating items for their own publications might create items manually. The approach taken is also dependent on what properties are desired (for example, is a text string for each author sufficient, or does the creator want each author to also have their own Wikidata item?). Overall, the past and ongoing efforts of the Wikidata and WikiCite communities can dramatically lower the amount of time and resources needed by publishers to add items describing their articles to Wikidata in a variety of ways, while publishers can offer the community institutional support that leads to consistency and longevity.

It might also be possible to enlist the aid of other services associated with journal metadata to reach publishers and editors. For example, a dialogue with Crossref and other DOI registration agencies could be initiated, with a request that when a DOI prefix is purchased the DOI registration agency include a link to a web page with an automated process to also create a Wikidata item using a DOI. This could be presented as an added value service to users of DOIs and other unique identifiers, with instructional videos and other information for those interested in the process and doing more. DOIs were first issued in 1998 and have been adopted by many journals and publishers as a way to provide a persistent and unique identifier for journal articles as well as graphs, data, illustrations, and other material in articles (DeRisi, Kennison, and Twyman 2003). The DOI project was launched by publishers through the Association of American Publishers, with Davidson and Douglas (1998, n.p.) writing that "publishers of the most costly scholarly journals . . . realize that long-term survival depends on their ability to market products successfully over the Internet." This suggests that demonstrating the value of linked data in increasing discoverability could motivate publishers to participate in its creation, as the survey results also suggest.

Although it would need to be more nuanced, the idea of demonstrating value through increased discoverability might also work for proposing that Wikidata items be created for authors when they register an ORCID iD. ORCID could provide a minimal Wikidata item as an extra service and (perhaps) offer authors the opportunity to add metadata about themselves via a Cradle (a template or form used to create Wikidata items), which would help alleviate privacy concerns as well as concerns about being accurate and sensitive in areas such as gender, ethnicity, sexuality, etc. Cradles are also customizable and could be created specifically for certain journals if desired. Furthermore, if authors listed their works in ORCID, and those works included a DOI, author name strings could likely be replaced with the Wikidata item for the author as part of a wholly or at least largely automated process. The Harvard "Online Author Questionnaire" (Harvard University 2018) has previously worked with publishers to gather information about authors (although not for entry in Wikidata) and could provide valuable insights into this process.

Ultimately, our survey results suggest that the greatest, or at least initial, barrier to journal editors and publishers creating Wikidata items for articles is a lack of awareness concerning what Wikidata is, what it can offer, and that it even exists. Editors and publishers recognize the value of metadata and support its creation but often view it through a cost-benefit lens. In order to dedicate resources such as staff time to the creation of Wikidata items, particularly the more time-intensive and sensitive areas such as items for authors, editors and publishers will need to be presented with concrete benefits that justify the cost. This is a far different ideological framework and motivation than the volunteer participants currently creating items because they believe in the mission of Wikidata and the intrinsic value of linked data, but one which does not automatically preclude collaboration. Therefore, a deliberate attempt to educate editors and publishers and demonstrate the value of creating Wikidata items is a logical first step. A case study with a small number of LIS journals could be prepared, one that documents the amount of time various steps require and how the resulting data can be used to support various goals, primarily traditional discoverability but also the possibility of discoverability through reuse of the data in knowledge graphs, lists, and other forms. From

this an educational document or presentation could be prepared and submitted to editors and publishers, in LIS and potentially other disciplines, to make a case for the creation of Wikidata items. While the materials would be designed primarily for journal editors and publishers, this approach could also be taken to university or library administrators by librarians to justify allotting staff time to the creation of items for faculty scholarship, theses and dissertations, and other creative and scholarly work associated with their institution. In all likelihood, some publishers and/or editors will be persuaded to create Wikidata items, and some will not (leaving plenty of work for volunteers to do in enhancing items), but it could be a first step towards ongoing institutional support for the creation of Wikidata items for journal articles across disciplines.

Appendices

- [Appendix 1: LIS Journal Publishers and Wikidata 2022-06 Survey](#).

CRediT Author Statement

Susan Radovsky, Harvard Library, Conceptualization and Methodology.
Eric Willey, Illinois State University, Conceptualization, Methodology, Writing – Original Draft, Writing – Review & Editing.

Acknowledgements

Competing Interests

Both authors are coordinators in the LD4 Wikidata Affinity Group.

Ethics and Consent

This study was evaluated and approved by the Illinois State University Office of Research Ethics and Compliance via Cayuse IRB before the survey was sent to any participants. The reference number for approval is IRB-2022-162.

References

Allison-Cassin, Stacy, and Dan Scott. 2018. "Wikidata: A Platform for Your Library's Linked Open Data." *Code4Lib Journal* 40. https://journal.code4lib.org/articles/13424.

Alves, Éder Porto Ferreira, Paul R. Burley, and João Alexandre Peschanski. 2021. "Structuring Bibliographic References: Taking the Journal *Anais do Museu Paulista* to Wikidata." In *Wikipedia and Academic Libraries: A Global Project*, edited by Laurie M. Bridges, Raymond Pun, and Robert A. Arteaga, 261–76. Maize Books. https://doi.org/10.3998/mpub.11778416.ch17.en.

Berg, Selinda, and Kristin Hoffmann. 2022. "Peer Reviewed LIS Journals." Last updated January 4, 2022. https://web.archive.org/web/20230512034025/https://library.usask.ca/ceblip/research-resources/peer-reviewed-journals.php.

Bianchini, Carlo. 2021. "Wikidata for *JLIS.It*. A New Step Forward Mapping Italian Library and Information Science Journals." *JLIS.It* 12 (1): 29–38. https://doi.org/10.4403/jlis.it-12680.

"Category:Discontinued Google services." n.d. Wikipedia. Accessed March 13, 2023. https://en.wikipedia.org/w/index.php?title=Category:Discontinued_Google_services&action=history. Archived at: https://perma.cc/TM26-6ZHQ.

Cobb, Simon. 2020. "Author Items in Wikidata." Paper presented at the WikiCite Virtual Conference, October 26, 2020. https://upload.wikimedia.org/wikipedia/commons/7/79/Author_items_in_Wikidata.pdf. Archived at: https://perma.cc/59L8-RYSA.

Davidson, Lloyd A., and Kimberly Douglas. 1998. "Digital Object Identifiers: Promise and Problems for Scholarly Publishing." *Journal of Electronic Publishing* 4 (2). https://doi.org/10.3998/3336451.0004.203.

DeRisi, Susanne, Rebecca Kennison, and Nick Twyman. 2003. "The What and Whys of DOIs." PLoS Biology 1 (2): e57. https://doi.org/10.1371/journal.pbio.0000057.

Flynn, Emily Alinder. 2013. "Open Access Metadata, Catalogers, and Vendors: The Future of Cataloging Records." *The Journal of Academic Librarianship* 39 (1): 29–31. https://doi.org/10.1016/j.acalib.2012.11.010.

Gregg, Will James, Christopher Erdmann, Laura A. D. Paglione, Juliane Schneider, and Clare Dean. 2019. "A Literature Review of Scholarly Communications Metadata." *Research Ideas and Outcomes* 5: e38698. https://doi.org/10.3897/rio.5.e38698.

Harvard University. 2023. "Online Author Questionnaire." https://projects.iq.harvard.edu/oaq. Archived at: https://perma.cc/X9N2-U4DF.

Kemp, Jennifer, Clare Dean, and John Chodacki. 2018. "Can Richer Metadata Rescue Research?" *The Serials Librarian* 74 (1–4): 207–11. https://doi.org/10.1080/0361526X.2018.1428483.

Kent, Will. 2019. "Why Is Wikidata Important to You?" *Wiki Education.* June 3, 2019. https://wikiedu.org/blog/2019/06/03/why-is-wikidata-important-to-you/. Archived at: https://perma.cc/97HS-FH49.

Lemus-Rojas, Mairelys, and Jere D. Odell. 2018. "Creating Structured Linked Data to Generate Scholarly Profiles: A Pilot Project Using Wikidata and Scholia." *Journal of Librarianship and Scholarly Communication* 6 (1): eP2272. https://doi.org/10.7710/2162-3309.2272.

Lemus-Rojas, Mairelys, Jere Odell, Lucille Frances Brys, and Mirian Ramirez Rojas. 2022. "Leveraging Wikidata to Build Scholarly Profiles as Service." *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 6 (3). https://doi.org/10.18357/kula.171.

McAndrew, Ewan, and Clea Strathmann. 2021. "Listeria: Create an Automatically Generated List for Wikipedia." https://www.ed.ac.uk/information-services/help-consultancy/is-skills/wikimedia/wikidata/listeria-create-an-automatically-generated-list-fo. Archived at: https://perma.cc/56MJ-QMRL.

Nielsen, Finn Årup, Daniel Mietchen, and Egon Willighagen. 2017. "Scholia, Scientometrics and Wikidata." In *The Semantic Web: ESWC 2017 Satellite Events. ESWC 2017*, edited by Eva Blomqvist, Katja Hose, Heiko Paulheim, Agnieszka Ławrynowicz, Fabio Ciravegna, Olaf Hartig. Springer, Cham. https://doi.org/10.1007/978-3-319-70407-4_36.

Odell, Jere, Mairelys Lemus-Rojas, and Lucille Brys. 2022. *Wikidata for Scholarly Communication Librarianship.* https://iu.pressbooks.pub/wikidatascholcomm/.

"Policy:Terms of Use/Frequently asked questions on paid contributions without disclosure." 2023. Wikidata. Last edited October 18, 2023. https://meta.wikimedia.org/wiki/Terms_of_use/FAQ_on_paid_contributions_without_disclosure.

"Scholia." n.d. Scholia. Accessed March 13, 2023. https://scholia.toolforge.org/. Archived at: https://perma.cc/9672-H5XQ.

Taraborelli, Dario, Jonathan Dugan, Lydia Pintscher, Daniel Mietchen, and Cameron Neylon. 2016. *WikiCite 2016 Report.* https://doi.org/10.6084/m9.figshare.4042530.

Thornton, Katherine, Kenneth Seals-Nutt, Marianne Van Remoortel, Julie M. Birkholz, and Pieterjan De Potter. 2023. "Linking Women Editors of Periodicals to the Wikidata Knowledge Graph." *Semantic Web* 14 (2): 443–55. https://doi.org/10.3233/SW-222845.

Vrandečić, Denny, Lydia Pintscher, and Markus Krötzsch. 2023. "Wikidata: The Making Of." In *WWW '23 Companion: Companion Proceedings of the ACM Web Conference 2023*, edited by Ying Ding, Jie Tang, Juan Sequeda, Lora Aroyo, Carlos Castillo, and Geert-Jan Houben, 615–24. https://doi.org/10.1145/3543873.3585579.

"WikiCite/Shared Citations." 2023. WikiCite. Last edited August 25, 2023. https://meta.wikimedia.org/wiki/WikiCite/Shared_Citations. Archived at: https://perma.cc/R6SY-NEVQ.

"Wikidata:Bots." 2023. Wikidata. Last edited February 12, 2023. https://www.wikidata.org/wiki/Wikidata:Bots. Archived at: https://perma.cc/8W9N-3CSF.

"Wikidata:Tools/Visualize data." 2023. Wikidata. Last edited October 21, 2023. https://www.wikidata.org/wiki/Wikidata:Tools/Visualize_data. Archived at: https://perma.cc/P5G8-3DXM.

"Wikidata:Wikicite." 2023. Wikidata. Last edited September 25, 2023. https://www.wikidata.org/wiki/Wikidata:WikiCite. Archived at: https://perma.cc/M3NQ-2ZDT.

"Wikidata:WikiProject LD4 Wikidata Affinity Group." 2023. Wikidata. Last edited November 27, 2023. https://www.wikidata.org/wiki/Wikidata:WikiProject_LD4_Wikidata_Affinity_Group. Archived at: https://perma.cc/X2BR-RS25.

"Wikidata:WikiProject PCC Wikidata Pilot/University of Washington." 2023. Wikidata. Last edited June 13, 2023. https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/University_of_Washington. Archived at: https://perma.cc/J8HQ-V7WS.

"Wikidata:WikiProject Periodicals." 2022. Wikidata. Last edited April 3, 2022. https://www.wikidata.org/wiki/Wikidata:WikiProject_Periodicals. Archived at: https://perma.cc/LMY9-TGQ7.

Wikidata Query Service. 2023. "Search for published in: British Journal of Educational Technology." https://w.wiki/76un.

Wilkins, A. Amina, Paul Whaley, Amanda S. Persad, Ingrid L. Druwe, Janice S. Lee, Michele M. Taylor, Andrew J. Shapiro, Natalie Blanton Southard, Courtney Lemeris, and Kristina A. Thayer. 2022. "Assessing Author

Willingness to Enter Study Information into Structured Data Templates as Part of the Manuscript Submission Process: A Pilot Study." *Heliyon* 8 (3): e09095. https://doi.org/10.1016/j.heliyon.2022. e09095.

Yon, Angela, and Eric Willey. 2021. "Learning from Each Other: Reciprocity in Description Between Wikipedians and Librarians." In *Wikipedia and Academic Libraries: A Global Project*, edited by Laurie M. Bridges, Raymond Pun, and Robert A. Arteaga, 288–301. Maize Books. https://doi.org/10.3998/ mpub.11778416.ch19.en.

Zuiderwijk, Anneke, Keith Jeffery, and Marijn Janssen. 2012. "The Potential of Metadata for Linked Open Data and Its Value for Users and Publishers." *JeDEM: eJournal of eDemocracy and Open Government* 4 (2): 222–44. https://doi.org/10.29379/jedem.v4i2.138.