
Promoting Vanderbilt Divinity Publications on Wikidata

by Charlotte Lew

GENESIS OF THE VANDERBILT WIKIDATA PROJECT

Vanderbilt's library staff started to work from home in mid-March 2020. Despite restricted access to library buildings, libraries continued to make services available to patrons. The majority of the staff had their daily operations reassigned to keep services vital. To grapple with the transition from onsite working to remote working, the library's leaders endeavored to create digital projects and invited the staff to collaborate on innovative proposals. The WikiCite project, also known as [VandyCite](#), was initiated in April 2020 by Clifford Anderson, the Associate University Librarian for Research and Digital Strategy, and was designed to add bibliographic data for Vanderbilt publications to Wikidata. Working on bibliographic data is second nature for most of the library staff, especially for catalogers, but editing bibliographic information on Wikidata required learning new protocols due to its unfamiliar data model and editing environment. Fourteen staff members, mostly without cataloging experience, participated in the project after watching a brief training video made by Cliff to learn the basics such as creating a personal account and adding items. The team met regularly for problem solving and strategic planning. These information exchange sessions not only sharpened our editing skills but also boosted work morale by strengthening our mental health and emotional resilience during the pandemic crisis. Every month, editing statistics were publicized in library town hall meetings as testimonials, celebrating our meaningful and measurable contributions.

LINKING DIVINITY PUBLICATIONS WITH WIKIDATA

In 2014, the Divinity Library collaborated with the Digital Scholarship Communication Team on an institutional repository (IR) project to promote open access for Divinity faculty publications (Anderson, Lew, and Warga 2016). This faculty publication archiving project collects comprehensive bibliographic data for each faculty member's scholarly output and populates the resulting data in Zotero. Thanks to the annually updated publication data housed in Zotero, the Divinity IR project lays the groundwork for the Divinity Wikidata project. The bibliographic data in Zotero can be automatically or manually added to Wikidata and is a game changer for productivity. By adding this data to Wikidata, the Divinity Library strives to elevate the scholarly reputations of faculty as well as staff and students.

More than 460 Divinity faculty journal articles with DOIs have been exported from Zotero. The bibliographic data linked to a DOI can be automatically extracted from [Crossref](#) by the deployment of Python scripts. Subsequently, the extracted data have been bulk uploaded to Wikidata. To facilitate this mass transposition, manually creating new items in advance for the journals that didn't exist in Wikidata enabled every article to be automatically matched up with the item for its journal by using the ISSN. After this dataset was manually filtered for error and enhanced if necessary, processing articles without DOIs was the next step.

Charlotte Lew is Coordinator of Digital Projects and Collections at Divinity Library, Vanderbilt University.

Steve Baskauf, the Data Science Specialist at the Heard Library, created VanderBot to automatically scrape data from Vanderbilt departmental websites and add researcher records to Wikidata. Owing to his accomplishment, all the Divinity faculty as well as the faculty in the Department of Religious Studies are included in Wikidata, providing a base for entering bibliographic data. The image below illustrates how VanderBot works.

VanderBot workflow

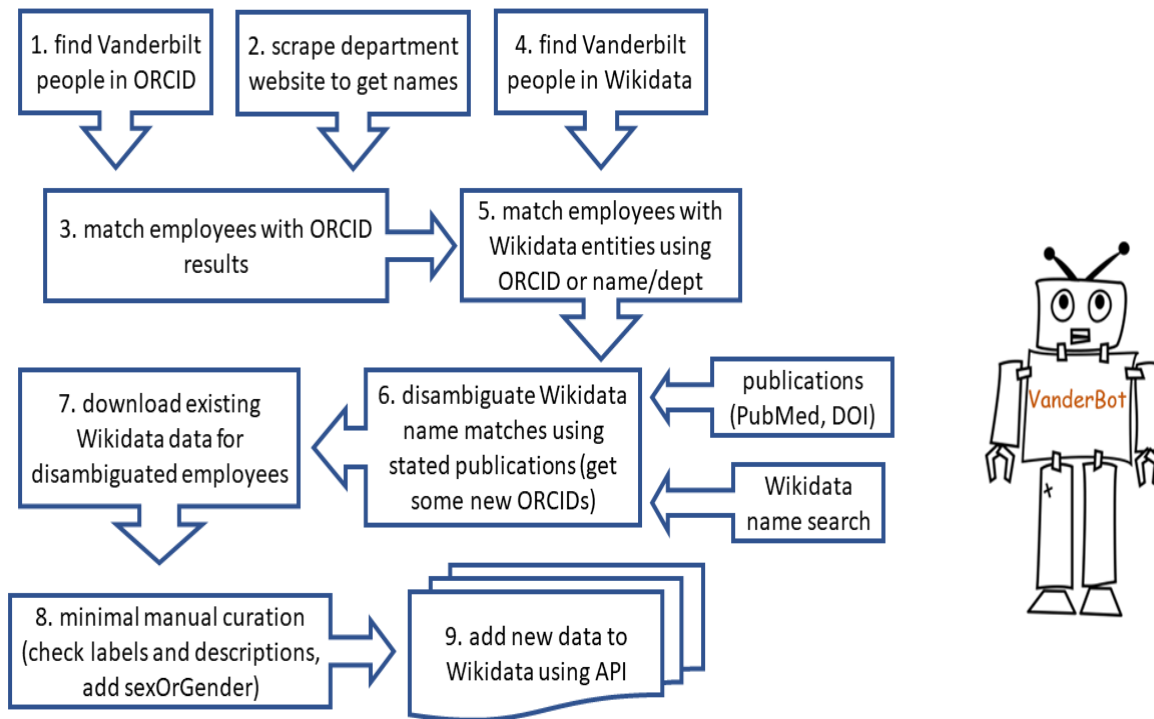



Image 1: VanderBot workflow created by Steve Baskauf

HOW DOES WIKIDATA WORK?

Wikidata organizes data in a specific structure, which consists of three primary parts: items, statements, and identifiers. Each item has a label, a description, and aliases if applicable. All the items have a unique identification with a Q followed by a number. Properties and values with qualifiers are used to build the statements. Identifiers are essential to Wikidata, as they provide crucial information for users to tell the difference between items with similar labels, such as people with the same name.

Choon-Leong Seow (Q15379207)

divinity school scholar  edit

Choon Leong Seow | CL Seow

▼ In more languages **aliases**

Language	Label	Description	Also known as
English	Choon-Leong Seow	divinity school scholar	Choon Leong Seow CL Seow

→ **Statements**

instance of **property** human **value**  edit

▶ 1 reference **qualifier**

+ add value

→ **Identifiers**

VIAF ID 7531915  edit

▶ 1 reference

+ add value

Image 2: Professor Choon-Leong Seow Wikidata page with indicators
<https://www.wikidata.org/wiki/Q15379207>

WHY SHOWCASE PUBLICATIONS ON WIKIDATA?

Wikidata is a central hub of open linked data for all the Wikimedia projects, and it is a multilingual platform for presenting and sharing information globally. Its impact on research, metadata production, and collection visibility is immense. As is well known, open access repositories in STEM disciplines are much more robust than those in the humanities and to some extent the social sciences. Therefore, increased access to religion and theology publications data on Wikidata is strategically important to researchers, instructors, students, and professionals in theological schools.

Wikidata is equipped with powerful tools that can be used to explore and improve Wikidata. All data in Wikidata can be queried with [SPARQL](#). Moreover, [QuickStatements](#) and [OpenRefine](#) can make uploading data easier and faster. The function of SPARQL queries can be extended to visualize the results. After the publications are added to Wikidata, researchers can view their publications via Scholia, an interface that supports Wikidata by organizing and visualizing the collected bibliographic data and other information using SPARQL queries. Besides listing publications, Scholia also offers visual scholarly profiles for authors and their affiliated institutions, and the included data visualizations can reveal new insights about the publications. The image below shows [Professor Laurel Schneider's](#) publication venue statistics, which are created by SPARQL on Scholia.

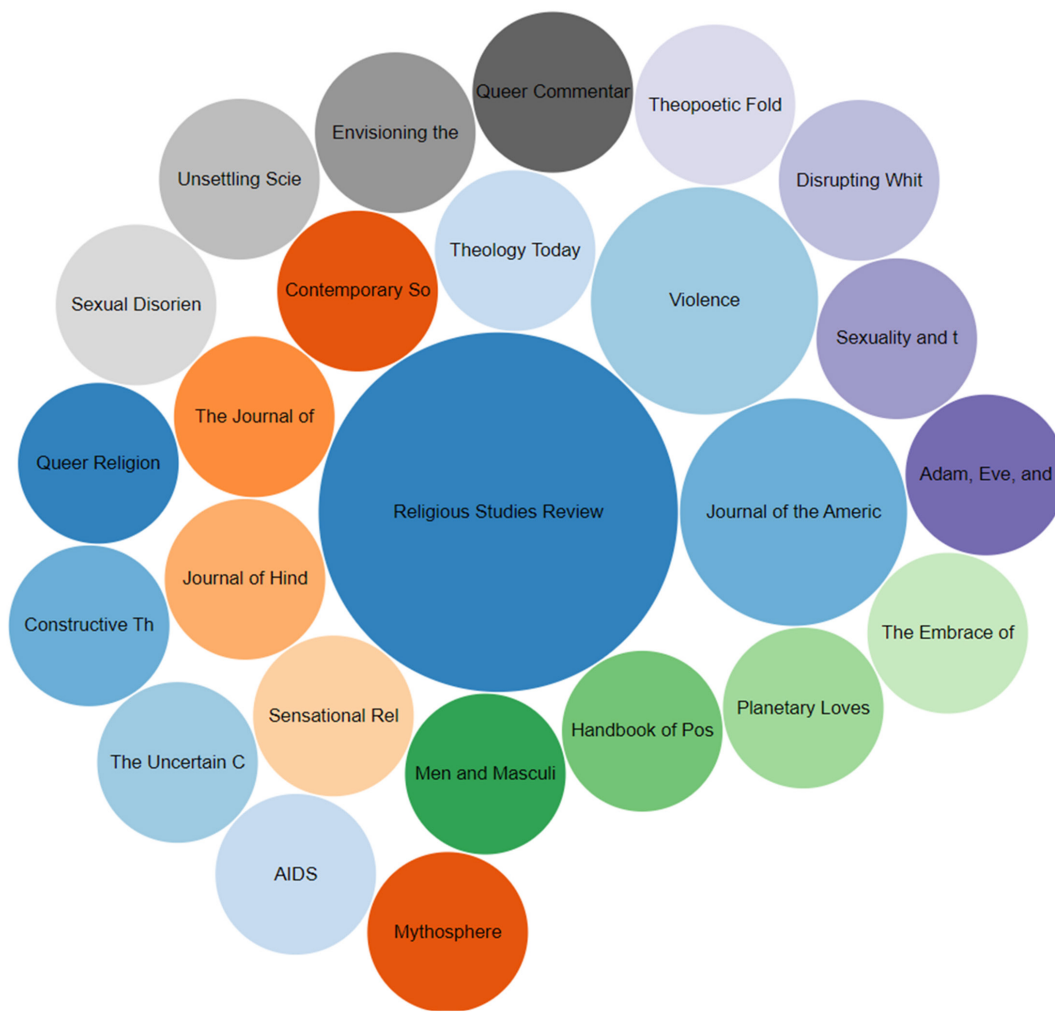


Image 3: The venue statistics of Professor Laurel Schneider's publications <https://scholia.toolforge.org/author/Q64009487>

CHALLENGES AND SOLUTIONS

When the Vanderbilt Wikidata project was first formed during the COVID lockdown, most of the team members taught themselves by watching the training videos. Nevertheless, the length of the videos can easily turn people away, and the needed information is often hard to locate in the videos. The most encountered issues were some basic questions about data model, structure, vocabularies, and completion in Wikidata. In the regular meetings, team members shared what they had learned from their experience. The following listed questions and solutions were spawned from the constructive conversations in the information exchange sessions. The purpose of sharing the following Q&As is to enlighten and prepare potential editors before they dive into productions.

[1. Where does the bibliographic data come from?](#)

As a part of the procedure of the Vanderbilt Institutional Repository Program, all the bibliographies of Divinity faculty publications are annually updated and harvested in Zotero. This low-hanging fruit empowers us to expedite the transposition of data from Zotero to Wikidata. To obtain the bibliographic data, the most effective route is to request a publication list from each faculty. Alternatively, one can search them from multiple resources, such as the library catalog, the Atla Religion Database, and the Index to Jewish Periodicals. Searching in Google Scholar can also be productive. Try searching for “faculty member name” “institution name.” Don’t forget to put the faculty member’s name in quotes. Using name variants may be needed for a more comprehensive search.

[2. What to do with duplicates?](#)

Wikidata doesn’t accommodate approximate string-matching for searching since all the entries are idiosyncratic, so it is important to conduct a thorough search before creating a new item. To Searching to avoid duplicates can be a time-consuming task because numerous items might share the same label, but with different descriptions. “The Prodigal Son”, for example, can be an opera by Benjamin Britten, a painting by Rembrandt, a novel by Hall Caine, or other various item types. When a duplicate is generated, there is no designated function key to delete it. Instead, the duplicated items should be merged into a single record. The operation of merging can be done by installing a merge gadget. The procedure is well-explained in [Help: Merge](#).

[3. How to make data reliable?](#)

Because Wikidata is open to anyone to edit, it gives rise to skepticism about the validity and reliability of the collected data. To minimize doubt, it is critical to add a reference whenever a statement is made. Providing references may be labor-intensive, but ensuring the presented information is accurate and factual should justify the time it takes. Fortunately, there is a way to speed up adding references. Wikidata offers a gadget called “current date” that adds today’s date automatically when the “retrieved” property is used in a reference. Enabling the different Wikidata gadgets is a boon to the editing experiences. The lessons of adding the various tools for editing are included in the training program, [Learn Wikidata](#), which will be introduced in last part of the article.

[*4. What to do with “item not found?”*](#)

When a statement is added, the warning of “item not found” appears when the entry for property or value is not an existing item in Wikidata. If this occurs while adding authors to a publication, it can be circumvented by using the property of “author name string” to replace the property of “author.” The downside of this practice is that the author can’t be hyperlinked to this publication. To ensure data are linkable, a new item must be created when there is an “item not found” warning prompt.

A good number of religious and theological publications are not published by major corporate publishers who have published data in Wikidata. As a result, numerous items related to the publications, such as publishers or journal titles, do not exist in Wikidata. It is strategically important to treat the creation of new items as a foundation on which data of the publications can be continually expanded, enriched, and rectified. Searching for the needed information to build the statements for a new item can be tedious and laborious. This can easily quench one’s editing enthusiasm, but there are several reference tools that can simplify the workload. For journal publications, Ulrichsweb, a global serials directory, brings together all the latest bibliographic and publisher information in one location. For books, Books in Print is a useful resource for retrieving information on global publishers. The drawback is that they both require subscriptions, but they are considerably worth the investment.

[*5. How to discover the proper property and value*](#)

Wikidata, a well-structured database, offers an intuitive workflow for editors to create or modify data. Workflow is arranged in a linear movement, starting from labeling items with descriptions to building statements and then providing identifiers. Each addition shows a drop-down of options for an editor to pick as the next property to fill. However, take heed of various types of property constraints. If an inappropriate value is entered in the property, a sign of a thunderbolt or exclamation mark will appear after the entry is published, but Wikidata identifies and explains the warning after the editor clicks on the sign. If an ISBN with 13 digits is not correctly separated by hyphens, a format constraint displays. For Rembrandt’s painting “[The Good Samaritan](#),” adding “religious art” as the value in the property “main subject” violates the type constraint. To rectify the violation, “religious art” should be listed in the “genre” property instead. The constraint list is lengthy and complicated. To learn more, please check out [Help: Property constraint portal](#).

Identifying key properties associated with particular item types is important. If a required key property doesn’t exist for the item type, [creating a property](#) is a practical solution. Don’t sweat over incomplete item statements. Making sure to include the core properties can be an easy start for beginners. Wikidata relies on global collaboration to enrich its data, so items that presently seem incomplete can be enhanced and augmented by other editors later.

COLLABORATE TO EVANGELIZE FOR WIKIDATA

Since Wikidata is a crowdsourced resource, it has a weakness of information unevenness. More publications from prestigious theological schools have been entered in Wikidata than those from

smaller institutions, so the participation of theological librarians from smaller institutions is essential to the future of this scholarly communications project. To lower the barriers so more people can join the editing force, the Vanderbilt Wikidata project team created a training tutorial called [Learn Wikidata](#). This peer-led course provides training that ranges from the basic skills of creating a user account and editing items to more advanced techniques like adding gadgets to make editing more efficient. The tutorial comprises twenty learning topics that are individually featured in brief animated videos. The program presents options of three languages: English, Spanish, and Chinese. French will be available soon. It also comes with closed captions to facilitate learning. To address the needs of librarians with various levels of experience with Wikimedia projects, Learn Wikidata offers a self-paced environment for users to customize their own learning pathways. The design of this training tool aims to alleviate frustration and to flatten the learning curve. Please give Learn Wikidata a try. Hopefully it can lead you to a pleasant editing journey.



Image 4: Homepage of Learn Wikidata
<https://www.learnwikidata.net/>

Wikidata provides a free platform for seminary libraries to publicize their institutional scholarship without having to invest in building an expensive and complicated scholarly communications infrastructure. To develop a community of practice around Wikidata, more theological librarians should be encouraged and empowered to become involved in these critically important projects. The Wikidata religion and theology community needs your partnerships to foster global scholarly communications that can elevate the reputation and increase the accessibility of your institution's publications. Let's collaborate and get stronger together!

WORK CITED

Anderson, Clifford Blake, Charlotte Lew, and Ed Warga. 2016. "Building Institutional Repositories in Theological Libraries." *ATLA Summary of Proceedings* 70: 153–63.