Opportunities for Academic and Research Libraries and Wikipedia

A Discussion Paper (Public Draft)
Presented at the IFLA World Library and Information Congress 2016 in Columbus, Ohio

IFLA

The International Federation of Library Associations and Institutions (IFLA) is the leading international body representing the interests of library and information services and their users. IFLA is an independent, non-governmental, not-for-profit organization with over 1300 members in nearly 140 countries.

We work to improve access to information and cultural heritage resources for the global community in this rapidly changing digital and print environment. Our key initiatives include access to digital content, international leadership, outreach, cultural heritage, and multilingualism.

In our professional programmes we build the capacity of our members, and set the professional agenda through development of guidelines, standards, publications and events around the world.

IFLA’s status as the global organisation for library and information services ensures that our voice is represented through formal relations with the United Nations and other organisations.

The Wikipedia Library

The Wikipedia Library (TWL) focuses on how to improve Wikipedia’s role in the research ecosystem. This started with helping Wikipedia’s volunteer editors access reliable sources to improve Wikipedia, and has since expanded to support other kinds of reference access and public engagement. As part of this mission, TWL helps knowledge professionals fruitfully share their collections with the public and engage with Wikipedia and other Wikimedia projects. The Wikipedia Library is funded by the non-profit Wikimedia Foundation.

Contributors

Stephan Bartholmei, Deutsche Nationalbibliothek
Rachel Franks, State Library of New South Wales
James Heilman, International Medical Translation Wiki Project
Mylee Joseph, State Library of New South Wales
Vicki McDonald, State Library of New South Wales
Anna Raunik, State Library of Queensland
Mia Ridge, British Library
Mark Robertson, Association of Research Libraries / Brock University Library

Supported by
Alex Stinson and Jake Orlowitz, The Wikipedia Library, Wikimedia Foundation
Julia Brungs, IFLA
Executive summary

This paper explores the potential for active collaborations between Wikipedia editors and academic and research libraries to advance the free flow of information which supporting sustainable development around the world. IFLA has the opportunity to support libraries, library staff and library associations by enabling their members to engage with Wikipedia: through networking, skill development, and showcasing examples of successful cultural and knowledge collaborations that can act as models for potential future initiatives. This paper explores several ripe areas of mutual benefit and collaboration around crowdsourcing and community engagement: writing and reference, technical tasks and linked open data, project coordination, high-speed publishing, learning communities, and the support of teaching, learning, and research. There is also consideration of the challenges of pursuing this vital work together.
Introduction

When the United Nations set a bold agenda for sustainable development, *Transforming our world: the 2030 Agenda for Sustainable Development*¹, the importance of access to information and knowledge to support sustainable development and improve people’s lives was captured in Target 16.10:

“Ensure public access to information and protect fundamental freedoms, in accordance with national legislation and international agreements” (UN Division for Sustainable Development, 2015).

In accordance with this mandate to make information publically accessible, the world’s research, university, public and special libraries continue to use physical and digital resources and skills to facilitate access to information from a wide range of sources to anticipate and meet the needs of their clients and the wider community; this position has been endorsed by more than 500 library and development associations and institutions in the *Lyon Declaration on access to information and development* (IFLA, 2014).

One of the multilingual tools widely available and freely accessible online is Wikipedia, a crowdsourced encyclopedia that is active in 282 languages with tens of thousands of editors and tens of millions of articles. The mission of the Wikimedia movement is directly compatible with that of libraries:

Wikimedia is a global movement whose mission is to bring free educational content to the world. Through various projects including Wikipedia and Wikimedia Commons, Wikimedia strives to bring about a world in which every single human being can freely share in the sum of all knowledge. (Cummings, 2016)

In order to reach these bold goals - to improve the lives of people through information - libraries will need to engage the widest possible range of online resources and engage with diverse partners and communities. Wikipedia plays an important part in this complex and multifaceted information environment. The collaboration plays advantage of both community’s core functions: as an encyclopedia, Wikipedia’s contributors rely on authoritative published sources to collect and curate quality knowledge summaries; meanwhile, libraries provide access to a range and depth of quality sources and research expertise, reducing Wikipedia editors’ reliance on less reliable publications and research methods.

This paper seeks to explore the opportunities libraries can pursue in collaboration with the Wikimedia community, the editors and developers who create and maintain the various projects, and resources of the Wikimedia Movement (that includes Wikipedia, Wikimedia chapters and

---

Opportunities for Academic and Research Libraries and Wikipedia: A Discussion paper

Thematic organizations around the world, Wikimedia Commons, Wikidata, GLAM-Wiki, and Wikisource). These collaborative opportunities offer a range of outcomes that could benefit academic and research libraries including crowdsourcing, community engagement, rapid models for online publishing, student writing assignments set in Wikipedia, and engagement with learning communities. The paper also identifies some of the challenges for libraries and library staff inherent in this type of online collaboration. With dialogue, knowledge exchange, and sufficient onboarding, libraries, library staff, and library associations can become savvy, literate contributors and collaborators together with the Wikipedia communities.

Opportunities for crowdsourcing and community engagement

By May 2016, the English Wikipedia had over 5 million articles, while the other language editions totalled more than 30 million articles. Wikipedia may be the most successful online crowdsourcing effort, and the largest curated collection of knowledge in human history. Lori Byrd Phillips argues (2014) that it is even more, emphasizing the collaborative approach of the Wikipedia community as a new and powerful form of "co-creation". The Wikimedia community negotiates how their sites will contribute to the goal of sharing knowledge, as well how to craft individual encyclopedic entries. The skills, knowledge and passion of the Wikimedia community span many areas including science and the humanities, history and biography, politics and georeferencing maps, linked open data, computer programming, photography, editing syntax, research, writing and translation.

Libraries and library staff can contribute to and partner with this community of content creators to develop information resources and improve access to information for citizens around the world. The core value of open access to information that drives the Wikimedia communities resonates with libraries and library professionals.

A number of best practices exist for engaging these communities: in 2015, Europeana published recommendations for managing its relationship with the Wikimedia community as part of implementing the Europeana 2015-2020 strategy (Wyatt, 2015), and OCLC has been active in providing information to libraries via webinars on how to work effectively with Wikipedia. Many examples of successful collaborations indicate a variety of opportunities for libraries and Wikimedia chapters and volunteers across a variety of types of tasks.

Writing and reference

Creating content on Wikipedia requires research across a range of published sources to ensure the information can be verified and shared with the public. Supporting research and writing are core priorities for academic and research libraries; disseminating information widely is a key motivation of the entire field.
Opportunities for Academic and Research Libraries and Wikipedia: A Discussion paper

- Wikipedia provides opportunity to reconcile physically disparate information resources in an online space and to collectively develop new knowledge and information. This can be seen in the Wikipedia article describing the *Journals of the First Fleet* which highlights the narratives around settlement of Australia by the British from the perspectives of those on board the ships of the First Fleet. This article, created by a variety of editors, draws on digitised resources from the State Library of New South Wales, transcriptions from the *Transcribe Bentham Project* at the University College of London and a wide variety of published sources.

- The digital editing platform provide by Wikipedia allows for editors to make contributions from any location, contributing a variety of perspectives that may be local and global. Contributions in more than 200 languages reflect both migration patterns and the ability of communities around the world to contribute regardless of location. Exposing collections and research materials in this ecosystem reaches experts around the world, and requires their knowledge to create.

- In many instances, Wikipedia has become a de facto public face of science and other fields of endeavour, acting as a pathfinder to secondary and primary sources.

- Further, there are currently 282 languages with active Wikipedias, with some articles translated into other languages and others written from scratch in each language. Though representing a wide and global perspective, many of the smaller languages do not have vibrant or active language communities, and many minority cultures that use a shared language are underrepresented -- thus offering opportunities for improved coverage and advocacy by stewards of these cultures and languages in libraries and archives.

Technical tasks and linked open data

Libraries face a number of technical challenges as they embrace emerging technologies, including digitization of material, linked data and mobile technology. The Wikimedia community has pioneered successful collaborative and open projects addressing some of these challenges:

- Geo-referencing information adds greatly to the value and usefulness of digitised maps and photographs. The Wikimedia Commons community has uploaded, classified and georeferenced thousands of the maps harvested from images digitised and released via the Flickr Commons in the *British Library's Mechanical Curator* project. These maps are now freely available for use and reuse via Wikimedia Commons, and many have been used to illustrate articles in Wikipedia. The Wikimaps Warper project has adapted the MapWarper developed for the NYPL to the Wikimedia environment allowing users to take scanned maps and add coordinates (Wikimaps, n.d).

- The importance of creating and publishing linked data is recognised widely in library circles, offering the potential to connect entities across many collections. The potential of Wikidata to draw linked data and linked data authorities together across the world’s languages and many different ontologies and taxonomies has enormous potential to support researchers around the world. For example, data from the *Disease Ontology*, an open source ontology designed to integrate biomedical data that is associated with human disease (Kibbe et. al, 2015), is contributed to Wikidata (Mitraka, et. al., 2015) where the
translations of disease names can be found, and the Gene Wiki initiative is exploring Wikidata as a semantic framework (Burgstaller-Muehlbacher, et. al, 2016).

- The recent WikiCite meeting developed a framework for migrating Wikipedia’s citation data into a linked open citation dataset deeply integrated with Wikidata and Wikipedia. Integrating more richly connected citation data alongside the structured data already in Wikidata would provide a web of connected data that could support innovative changes in library discovery tools.
- Linked open data has enormous potential for discovery across diverse collections, and Wikidata offers semantic enrichment of library records. For example, the National Broadcaster of Finland, YLE, has begun using Wikidata concepts to tag content in their digital archive. The Wikidata Query Service provides a SPARQL endpoint to query against the Wikidata data set and there are a number of other ways to create linked data using Wikidata.

Providing access to information in regional and remote areas and to some of the world’s most vulnerable communities is also a priority:

- In some countries, Wikipedia is accessible to users on mobile devices for free via the Wikipedia Zero project.
- It is also possible to download Wikipedia for offline access via projects like Kiwix and the popular offline Wikipedia Medicine app.
- There are additional ongoing efforts in collaboration with Internet-in-a-Box to create travelling offline distribution systems. Plans exist to trial a number of these units, loaded with Wikipedia content, in medical libraries in the developing world in the second half of 2016. Other projects in the past, spearheaded by research libraries, have provided similar offline Wikipedia access: for example, the Egranary project supported by the University of North Carolina, Chapel Hill.
- Wikipedia is included in the Ideas Box portable media center toolkits deployed in refugee camps.

Project coordination

Libraries around the world are responding to a shift in information-seeking behaviours, including an increasing reliance on mobile technology, and easy access to Wikipedia and other free online information sources.

- It may prove valuable for library staff to develop collaborative partnerships with the communities creating some of these online information sources. For example in Armenia the local Wikimedia chapter is very active in training editors and holds regular wiki camps and conferences.
- Many Wikimedia communities stage edit-a-thon events (e.g. Wikimedia Mexico’s recent world-record-breaking event). Libraries, museums and archives have hosted themed edit-a-thon events on specialist topics including women in science and the Africa Collection Editathon at the Library of Congress.
The volunteer base and community outreach of Wikimedia projects can engage people from around the world in creating content (e.g. Europeana's 280 project is a multilingual editing blitz celebrating more than 80 international artworks). Wikimedia also runs a number of photography competitions on different themes where many high quality freely licensed images are created, these include Picture of the Year, Wiki Loves Earth, Wiki Loves Monuments, Wiki Loves Africa and Picture of the Day.

There are opportunities for libraries with key subject strengths to engage on a relevant WikiProject, an organized group of participants on a Wikimedia project established to achieve specific editing goals or to achieve goals relating to a specific field of knowledge.

Wikimedia chapters initiate and partner on significant research and publication projects related to open knowledge, copyright policy and other projects relevant to libraries (e.g. Wikimedia Deutschland's participation in the development of Practical framework for Open Educational Resources (OER) and Open Content - A practical guide for the use of Creative Commons licenses with government and other organisations).

High-speed publishing

One of the strengths of Wikipedia is the speed with which information can be published and updated, allowing rapid responses to urgent and evolving situations like medical epidemics and natural disasters with information and translations contributed from many editors. Library staff may be vital contributors and editors in these situations by updating and citing information; they also contribute by providing access to quality resources other editors can use about current events (including politics, legal events, and scientific discoveries). There are a range of ways libraries can facilitate these editing activities:

- An important role that library staff can play is to support Wikipedia editors who need to verify and confirm data they use through library collections and online resources (e.g. the National Library of Israel Reference Desk project on Hebrew Wikipedia and the “Ask a Librarian” links that WikiProject Australia has added on relevant article talk pages).
- Libraries can provide a citation option for books and newspaper articles listed in their catalogues in Wikipedia's format and markup (e.g. National Library of Australia's Trove catalogue).
- It is possible to integrate Wikipedia results into some library discovery systems. For example, at the State Library of Queensland someone searching the catalogue for Peter Finch is presented with library holdings as well as content from the the Wikipedia article, and someone looking for a digitised newspaper in the National Library of Australia’s Trove catalogue will find the entry enhanced with content from Wikipedia.
- Information about natural disasters and current events can take mere hours to reach the mass media and be summarized on Wikipedia, but it further develops over months to appear in mainstream periodicals and peer-reviewed journals. Examples of Wikipedia articles that quickly captured a critical incident and continue to evolve as research and investigations unfolded can be seen in the 2004 Indian Ocean earthquake and tsunami article on English and the same on other languages (e.g. Bahasa Indonesia Gempa bumi dan tsunami Samudra Hindia 2004 and Vietnamese Động đất và sóng thần Ấn Độ Dương 2004).
Wikipedia editors also collaborate with groups in crisis situations, including Wiki Project Med Foundation’s work with Translators Without Borders to improve the quality of health related information on Wikipedia’s less-developed languages. For example, during the Ebola crisis, content on Wikipedia was developed in more than 100 languages and Wikipedia was the single most used source of online information in the most affected countries (Cohen, 2014).

Open access publishing frequently draws material from a range of sources, including Wikipedia articles, repurposing it in apps and on websites and in other types of publications. (Weller, 2011).

Learning communities

Wikimedia is a global learning community. It is not only open to all, but it is continually evolving as it incorporates new models of content production and curation. Wikimedia is also a constant forum for sharing among its participants, in the effort to improve the capacity of its thousands of volunteers, and to welcome and equip new contributors.

- This community of volunteer editors takes many forms, from individuals developing skills and knowledge as they contribute, to supporting new editors via online projects like the Tea House, WikiProjects and peer to peer coaching.
- The community is assisted in learning through Wikimedia Foundation supported initiatives including the Evaluation Portal sharing resources to better understand Wikimedia programs and a Learning Pattern Library for WikiProject leaders.
- The community also holds conferences such as GLAM-Wiki where volunteers who are working on projects that involve galleries, libraries, archives and museums (GLAM) can come together and share their experience and expertise. Other opportunities to engage with the community include Wikimania and Wikimedia Conference, discussions for special interest groups, edit-a-thon events and collaboration with Wikimedia chapters.
- There are many opportunities for libraries reaching out to their local Wikimedian communities through staff as content creators, institutions contributing public domain images, collaborative partnerships, hosting edit-a-thon events and liaising with local Wikimedia chapters.
- Opportunities include building capability and partnerships such as developing staff editing skills. The future education of library professionals could include contributing to Wikipedia and a deeper understanding of the role of Wikipedia in the information ecosystem.
- A particularly effective initiative involves libraries hosting Wikipedian In Residence programs (Gray, 2013) where a skilled editor is available to provide expertise, assist in management of collaborative projects and identify potential engagement with Wikipedia and Wikimedia projects through content donations, editing and coaching staff in developing skills.

Teaching, learning and research

Wikipedia as a digital platform, a range of technical projects, and a community of editors offers many opportunities to academic and research libraries. In an era of diminishing resources libraries
can align the benefits of contributing to Wikipedia with their core missions and objectives to meet the needs of library clients in multiple ways: information literacy, exposure of library collections, responding to client information seeking, leveraging library authority data, and addressing the needs of digital humanities researchers and digital scholars.

- Library professionals frequently teach information literacy skills to library clients and students. Wikipedia can be used as a tool for teaching information literacy. Familiarity with the features of Wikipedia pages that allow librarians to teach key digital information literacy strategies that patrons can apply to other digital content, such as checking the editing history, locate the sources cited and investigate the talk pages for discussions about the content and bias in an article (Bishop, 2015).

- Research indicates many higher education students use Wikipedia as part of their information seeking process for course-related research, with studies citing its "coverage, currency, convenience, and comprehensibility" (Head and Eisenberg, 2010). In many ways Wikipedia articles act as highly accessible pathfinders to information sources. Teaching students to connect the material on Wikipedia with other research materials teaches critical research skills.

- Around the world contributing to Wikipedia has been incorporated into teaching in higher education, giving students valuable experience "collaborating in a diverse environment, fostering critical thinking about evidence, effectively communicating information to a broad public audience and appreciating the role of access to information in a democratic society" (Infield and Adams, 2013). Support for this kind of “Wikipedia Education Assignment” is available throughout the world. The Wikipedia community has developed guidelines for designing student assignments that contribute to the encyclopedia while also complying with their policies and guidelines.

- Using edit-a-thons and other skill building activities, such as public participation in scanning and transcription of collection materials, to provide hands-on programming around library collections, engages scholars and researchers from the broader public with relevant materials.

- WikiProjects can be engaged in their focus areas to leverage Wikipedia for exposing and enhancing the discovery of specialized collections.

- Efforts are being made to ensure that citations lead back to library collections and open access research materials; for example a local history edit-a-thon that highlighted material in the UMass Lowell collections (Diaz, 2016).

- Increasingly, academic libraries are expanding services to include access to technology and support for the creation and curation of digital material for research purposes (e.g. digital humanities labs). Many Wikimedia projects can provide platforms or data that can be included in digital scholarship across the humanities, in geographic information systems, and for text analysis and data visualisation.

**Challenges for library staff**

Contributing to Wikipedia also poses some unique challenges for libraries and library professionals:
One of the core principles of Wikipedia is maintaining a “neutral point of view (NPOV)” and the precautions around “conflict of interest” editing. This hinders library staff or Wikipedian-In-Residences from editing articles about their own institutions. Despite these rules, dozens of institutions have worked within this constraint to develop rich content from their collections and about their organization’s history and people.

Institutions may need to institute and official staff editing policy (e.g. State Library of NSW Staff editing policy). This is particularly important as the editing process can occasionally be confrontational and having an agreed consistent policy and framework supports staff in interactions with Wikipedia editors and reviewers.

Institutions can encounter problems if volunteer editors are not aware of their presence, identity, or aims. An institutional project page is highly recommended as a means of communicating with the millions of other Wikimedia editors the intent of the library’s participation, planned activities and work in progress, and compliance with the policies and guidelines of Wikipedia (e.g. British Library Project Page).

Wikipedia has its own intensive peer review processes (Feature Article, Good Article and Quality ratings), editing policies, manual of style and templates which are not generally well known outside of the expert editing community. Library staff can learn to operate within these policies and constraints once they are aware of them, but navigating them is a barrier to entry for newcomers. Building this knowledge is much easier with a Wikimedian-In-Residence to coach staff in editing. Fortunately, Wikipedia’s policies have deep similarities to expectations found in academic writing, so this is only an onboarding challenge.

Wikipedia’s commitment to Creative Commons-licensed or public domain media is firm and at times more radical than libraries are familiar with. Donations to Wikimedia Commons, the free media repository associated with Wikipedia, are appreciated, and can be used to illustrate Wikipedia. But it is vital that library websites clearly and unambiguously communicate the copyright status of a text, image, or video in order to engage Wikimedia editors without having their donations removed or deleted. (e.g. compare a typical library catalogue record with Wikimedia Commons).

Libraries may employ unstable or impermanent links to their content. Where libraries deploy permalinks, Wikipedia can demonstrably increase the traffic to a library website or online collection.

Libraries may each use different, unique, or select identifiers and authority control methods for disambiguation. Increasingly, Wikidata is becoming an important tool for synchronizing across identifiers like Virtual International Authority File (VIAF) and ORCID identifiers. Integrating awareness of Wikidata and its uses for enhancing metadata and linked open data will help advance a more interconnected research web.

Libraries are often required to report on return on investment and demonstrate success. The development of reliable metrics and accurate reporting from Wikipedia and WikiProjects is key to demonstrating the value and reach of these activities.
Conclusion

It is evident through a decade of shared projects that libraries and Wikipedia editors have enormous potential for collaboration and the exchange of skills and expertise. A great deal of this potential is dependent on building connections between libraries, library associations, library staff and Wikipedia editors, Wikimedia chapters, GLAM-Wiki networks, and as yet uninvolved individual volunteers. IFLA and the Wikipedia Library project can be instrumental in building those bridges and facilitating the connections and networks.

Together Wikipedia can foster active outreach to libraries, help train library professionals to understand Wikipedia's policies and guidelines, support new contributors in acquiring technical editing skills, develop information literacy required to assess the quality of Wikipedia's articles, broaden awareness of WikiProjects and initiatives including participation in GLAM-Wiki events, encourage Wikipedian-In-Residence placements to involve staff in coordinated writing efforts and bulk upload of images, promote Wikimedia activities to local communities and faculty, create toolkits about re-using Wikipedia's content and metadata, and disseminate strategies for collecting metrics that demonstrate the research impact Wikipedia can provide.

Equally, the Wikimedia communities around the world need libraries and librarians. Wikimedians need access to quality sources to continue building the encyclopedias and other online tools, the expertise of information professionals around the world who can recommend and advocate for a diversity of types of knowledge from many languages and cultures, institutional libraries that are willing to take on Wikipedian-In-Residence initiatives as expanded partnership and skill development efforts, institutions contributing back into the WikiData metadata ecosystem, and librarians getting involved in organizing with and influencing Wikimedia's own chapters and thematic groups.

There are cultural, political, technical, and administrative challenges to overcome in advancing these diverse efforts broadly and at scale. In context, however, the opportunities are great, and the alignment of interests clear— and with existing case studies proving that value, it's the time to invest in this mutually beneficial alliance between Wikipedia and its volunteer communities and Library institutions and librarians. Many more experiments will yield best practices, and in-depth conversations will lower barriers that stand in the way of tremendously promising collaboration between these two distinctly important spheres of influence, each in their own way stewards of human culture and knowledge.
References


Editing Wikipedia Guide
https://upload.wikimedia.org/wikipedia/commons/1/18/Editing_Wikipedia_brochure_EN.pdf


Reisz, Matthew (14 August 2014). "Wikimania: student medics get credit for website manner". *Times Higher Education*.


Teigen, Sarah (October–November 2012). "Medical translations for minority languages". *Multilingual*.


