Marking up NSW: Wikipedia, newspapers and the State Library


Simon Cootes
A/g Collection Development Librarian
State Library of New South Wales
simon.cootes@sl.nsw.gov.au

Abstract:
This paper discusses the State Library of NSW’s project to create content on Wikipedia linking back to the digitised newspapers on Trove. The paper discusses the issues encountered in working with Wikipedia and the factors that made this project a success. One significant factor was the collaborative partnerships with other institutions that allowed the State Library to build on the expertise of other groups.
Introduction

The State Library of New South Wales (SLNSW) aims to be an innovative gateway to information and cultural enrichment. In 2012, SLNSW established a Memorandum of Understanding with the National Library of Australia (NLA) to digitise and make public selected NSW newspaper titles. Under this agreement, the two libraries are working collaboratively to digitise and provide online access to six million historic pages via Trove. Trove is a discovery tool developed by NLA that focuses on resources relating to Australia and Australians. Since 2008, it has made digitised Australian newspaper available to the public.

In 2012, SLNSW also began a project to explore new ways of providing service delivery online, including through Wikipedia. The project aimed to create Wikipedia records for NSW newspapers as they were digitised and made available on Trove. This paper discusses the nature of the project work and the partnerships in which SLNSW engaged with various organisations. These partnerships were integral to the success of the project. The paper discusses the benefits of creating and editing Wikipedia content, and presents the impact of these Wikipedia records on traffic to SLNSW’s two catalogues and to Trove.

The Innovation Project

In 2008, SLNSW established the eRecords Project, to make the entire collection available via digital bibliographic records, promote discovery and transform business processes as a result. To assist with the business transformation process, SLNSW established the Innovation Project in 2012 to pilot new services in the online world. It aims to extend SLNSW’s existing use of social media tools, with a view to developing a model to support and enable service delivery innovation. The new services are delivered from both the SLNSW website and in the online spaces frequented by the Library’s clients. To identify these spaces, SLNSW examined rankings for social media websites as well as the practices of other libraries. The Innovation Project established a number of work packages to explore different social media platforms and develop staff skills in using these platforms. The end goal of the Innovation Project is to have these platforms and work processes become ‘business as normal’ for SLNSW. The purposes of the Innovation Project are to:

- raise awareness of SLNSW and what it has to offer in the digital world
- actively seek out new and existing clients in the online communities they inhabit
- provide more staff with the opportunity to use social media tools to deliver services
- help develop and embed skills to increase our agility in responding to client needs in the online world.

One of the tools explored as part of the Innovation Project was Wikipedia. Although not seen to be a social media platform by the Wikipedia community, this crowd-sourced platform was held to be of immense potential to libraries. Wikipedia is the sixth most visited website in the world as listed by Alexa’s Top Sites website (Alexa n.d.), ranks highly in search algorithms and includes a number of different language editions, all drawing from the Wikimedia Commons image bank. In a recent Pew Internet survey
(Purcell et al. 2012, ‘The Changing Definition of Research’, para 3), teachers indicated that Wikipedia was the second most frequently used resource by students.

By using Wikipedia, SLNSW aimed to enhance the discoverability of collection materials and content through editing existing Wikipedia articles and creating new articles with links back to the SLNSW catalogues and to Trove. The other goal to be achieved by working with Wikipedia was an increase in client and public library engagement through the encouragement of content editing on Wikipedia. The Wikipedia Project consisted of two separate strands of work. The first strand, which is still ongoing, intends to create a Wikipedia record for every digitised NSW newspaper on Trove. The second strand was aimed at creating or editing records connected to the 100 collection treasures displayed as part of the Mitchell Library’s centenary.

Libraries and Wikipedia

Zlatos (2012, p. 339) notes that while reference librarians rarely begin a reference search with Wikipedia, many are now contributing to Wikipedia records in order to benefit from Wikipedia’s usage levels or as a way to ‘promote the credibility of librarians’. Badke (2008) also argues that academics and librarians should be contributing to Wikipedia. Librarians, like academics, are members of a trusted profession, and by contributing to Wikipedia, can aid in the promotion of a more scholarly environment within Wikipedia. Wikipedia is one of the world’s most visited websites, and through working with Wikipedia, libraries can promote the awareness of collections that may otherwise be little known or utilised. As Lally and Dunford (2007) have pointed out, libraries need to look for new ways to reach users where they are seeking information. By taking their content to Wikipedia, libraries are going to where their clients are active, instead of waiting for them to come to the library space. This is particularly true for digitised collections due to the ease with which this content can be linked to from Wikipedia (Pressley and McCallum 2008).

Before starting this project, SLNSW considered the benefits and risks of working with Wikipedia. A social media brief was prepared by SLNSW outlining the project’s aims, target audience, risk factors and the measures of success. In preparing for the project a number of conversations were also had with internal stakeholders, the public library network, other National and State Libraries of Australasia (NSLA) institutions and Wikimedia Australia. These conversations covered a number of areas including the experiences of other institutions in working with Wikipedia. The State Library of Queensland (SLQ) has previously worked with Wikimedia Australia and the Queensland public library network on a similar project to add authoritative content to Wikipedia ('Wikipedia:GLAM/State Library of Queensland', n.d.). SLQ’s experience was used to help inform the decisions made at SLNSW and a training manual developed by SLQ (2013) was also used in training SLNSW staff working on Wikipedia.

The experiences of libraries other than SLQ were also considered in planning the Wikipedia Project. Much of the work on similar projects has been undertaken by academic libraries. In 2010, Ball State University ran a project to enhance discoverability of their digitised assets through Wikipedia (Szajewski 2013). The University has seen page views of their digitised Hague Sheet Music Collection roughly triple since placing links and
information in Wikipedia. This is an example of the benefit of using Wikipedia to promote digital collections.

However, libraries have encountered problems in working on Wikipedia. One such obstacle is Wikipedia’s insistence on individual accounts. Organisational accounts are not permitted and this has led to records being deleted and accounts removed by Wikipedia, as happened with Wake Forest University (Pressley and McCallum 2008). Staff at the University’s Z. Smith Reynolds Library initially experimented with Wikipedia and subsequently had their articles deleted and their account deactivated. Only after researching the issue did they discover that Wikipedia does not allow for corporate accounts, and that their article could be taken as an advertisement for their collections. Another example occurred in 2006, when the University of Washington ran a project to insert links to their digitised collections in Wikipedia (Lally and Dunford 2007). The University encountered difficulties when staff, not using registered accounts, rapidly added links to articles. This activity is a red flag for Wikipedians who watch for vandalism. In both of these examples, however, the libraries concerned learnt from their mistakes and went on to contribute successfully to Wikipedia. To avoid these problems, SLNSW project staff created individual accounts containing clear statements that they were working on Wikipedia in a professional capacity as SLNSW staff.

**The Wikipedia Project**

In planning the Wikipedia Project, a decision was made by the project leaders to create only ‘start’ class records for each newspaper where a Wikipedia record did not already exist. A start class record provides a basic level of information but is in need of further development (‘Template:Grading scheme’ n.d.). This decision was made as the sheer number of titles requiring records created dictated that only a minimum level of research could be done for each title. Seventy-nine newspapers were initially in scope for the project. It was decided that there was too much work involved in researching every newspaper in depth for the level of detail required for a high-level record. Making this decision also allowed SLNSW to engage in partnership with the NSW public library network. By creating start class records, SLNSW was able to provide a basic level of information across a wide range of titles. These records could then be easily edited by the NSW public library network or anyone with an interest in NSW newspapers without needing to create a record from scratch. This is just one way in which SLNSW is able to make use of its resources to support the public library network.

Although the newspaper records were start class records, records created for the 100 collection treasures were produced as high-level records. These included the *1813 Crossing of the Blue Mountains* (n.d.) and *Governor Davey’s Proclamation* (n.d.).

For each of the newspaper records, a minimum level of content was included to cover basic details. This level was determined in consultation with Wikimedia Australia and experienced cataloguing staff. The included level of content was deemed to be the minimum required to meet Wikipedia’s notability requirements for records (‘Wikipedia: Notability’ n.d.). The content included the following details:

- Years of publication
Wikipedia policies dictate that information within records be encyclopaedic, written from a neutral point of view and be verifiable. Assertions must be qualified with references, and it is this insistence on authoritative referencing that has aided Wikipedia in gaining some acceptance amongst academic and educational circles (Lamb and Johnson 2013). This is part of what Wikipedia refers to as its “Five Pillars” (‘Wikipedia: Five Pillars’ n.d.). These policies, and Wikipedia’s notability policy for articles, also meant that thorough research for the articles was imperative as part of our project. For an article or list to be created, the topic must meet Wikipedia’s notability requirements. Wikipedia’s general requirement for notability states that a topic will be considered notable “if a topic has received significant coverage in reliable sources that are independent of the subject” (‘Wikipedia: Notability’, n.d.). This coverage must be in the form of verifiable independent sources.

For these reasons, the traditional information management skills of the project staff proved invaluable. In creating the newspaper records, project staff consulted a wide range of resources to establish historical facts about the newspapers’ publication histories. As many of the newspapers are otherwise obscure titles for regional publications begun in the nineteenth century, these facts could be difficult to locate and were rarely found online. Even though the articles were start class records, research could often prove difficult.

For the second project strand, dealing with collection treasures, the research requirements for detailed records were extensive. The research for the newspaper group was aided by the compilation of a relevant bibliography by the project coordinator shortly after the start of the project. A list of sources was compiled, including online and print resources that were readily available to project staff. The works of Rod Kirkpatrick, who has published a great deal of research on newspapers in NSW, were used extensively. Without these works, much of the basic history of country newspapers would have been difficult to locate, and hence notability would have difficult to establish. Other useful sources for the early history of newspapers included the Australian Dictionary of Biography and the newspapers themselves.

One resource often used as a citation for publication dates in Wikipedia records was the SLNSW catalogue for published material. As the project progressed, however, it became clear that the catalogue information, in several instances, was not as accurate as was expected. On separate occasions, it was discovered that information from the catalogue conflicted with information cited in other reference sources. One example of this was seen in the Southern Morning Herald. A Wikipedia record for this was created using information from the catalogue record. SLNSW was subsequently contacted by a public librarian who had attempted to add information to the record but realised that the existing information conflicted with her own research. With further investigation, it was established that the catalogue record was in error.
Once the records were created, SLNSW communicated with the public library network via email lists and through library training sessions to make libraries aware of the records' existence. Public libraries were then able to edit the records and add their own content to Wikipedia. Public libraries are best situated to add this content due to their excellent connections with local studies groups. Local history groups can make use of local studies resources to provide a greater amount of relevant content for the records.

It was initially expected that the basic nature of the newspaper records would mean that the amount of time needed to be invested in research and record creation would be minimal, and that this would decrease as the work involved became business as usual for SLNSW. This assumption was also based on the premise that the work would be similar to that of a cataloguer for published materials. However, as the project progressed, it was discovered that more research time was required than expected, as the work involved was closer to that of a curator who may be involved in describing archival or unpublished materials. The average time to create a Wikipedia start class record for a newspaper was 1.5 hours. Since the end of the initial project, this has decreased to approximately one hour. This included the time required to research and write the article, as well as upload an image. In contrast, a single detailed article on the 1813 Crossing of the Blue Mountains required a total of 10 hours for researching and writing the content.

A total of fourteen staff worked on the two strands of the project. The project staff work in a range of different positions across SLNSW, including library assistants, library technicians and librarians. Senior librarians acted as project coordinators. As one of the project aims was to extend Wikipedia skills to all staff levels, no distinction was made in distinguishing the nature of work performed by each staff member. Staff were selected for the project through an invitation to submit an expression of interest for the project. The project group predominantly worked together as a group during weekly 3-hour sessions. Some sessions were dedicated to training staff in particular aspects of Wikipedia, such as the uploading of images. Staff were released for the project following negotiations with their team leaders regarding workloads and existing commitments.

**Partnerships**

The commitment to create a record for each of these newspapers was made in accordance with a Memorandum of Understanding established with NLA. In 2012, SLNSW established a Memorandum of Understanding with the NLA to digitise and make public selected NSW newspaper titles. Under this agreement, the two libraries will work collaboratively over the next four years to digitise and provide online access to six million historic pages via Trove.

The partnership has already resulted in a large number of historic NSW Newspapers and Government Gazettes being digitised. In its first year alone, a total of 90 new titles were digitised, with 2.4 million pages scanned. Through this partnership, SLNSW agreed to create Wikipedia records for each newspaper title, and these would contain direct links back to the newspaper on Trove. The Wikipedia records were to be created in such a way that the introduction from each record could then be ingested into Trove to provide a basic level of description for Trove users looking at the newspapers. SLNSW and NLA have collaborated on what level of information will go into these introductions. Each month,
SLNSW submits a list of records that have been created or edited to NLA, where further checking is done. Any issues with the records are then discussed by the two libraries before the data is pulled from Wikipedia for inclusion in Trove. In addition to the partnership with NLA, SLNSW has also collaborated with Wikimedia Australia and the NSW public library network for this project.

In order to avoid the problems other libraries have encountered with Wikipedia, Wikimedia Australia was approached prior to the commencement of the project. Wikipedia has for a number of years sought to include more input from libraries in their records, through activities such as 'editathons' and the Wikipedia Loves Libraries series of events in the US (SinhaRoy 2011). Wikimedia Australia was approached by SLNSW, as it was known to have worked successfully with SLQ on a similar project. SLNSW aimed to enlist the expertise of Wikimedia Australia’s staff in scoping the project and for the training and support of project staff. Due to the fact that Wikimedia Australia’s trainers are volunteers, discussions were required to find a level of training which suited the trainers’ availability and met SLNSW’s requirements.

Wikimedia Australia provided several experts who conducted training sessions for staff. This training took place initially as one-day workshops for project staff prior to the commencement of the project. These provided staff with the information necessary for creating records and participating within the Wikimedia community. They helped project staff to prepare for the work they would be engaging in and allowed them to make use of the knowledge of experienced Wikipedians. The lessons were planned by Wikimedia Australia to allow SLNSW staff to begin editing as soon as possible. Wikipedia policies and practices were discussed as staff encountered them in their work, and this helped to keep staff engaged in the work.

The training sessions were highly successful, as the participants were already committed to the project and its success. Staff were selected for the project through an application process rather than by nominating staff directly. This helped ensure staff were committed to the project, as all involved had volunteered to work on the project. As a result of this established commitment, the trainers were able to focus on how things should be done, rather than on why things should be done. The existing reference and research skills of the librarians were acknowledged by the trainers as being valuable to the process and relevant to the work of Wikipedia (‘Wikipedia talk: GLAM/State Library of New South Wales/November 2012’ 2012).

The success of the training was also in large part due to the extensive planning beforehand. This planning included the scheduling of the sessions and, most importantly, the allocation of target articles to be edited by the staff. Wikipedia accounts for the staff were created prior to the training, which meant that little time was wasted on administrative functions. Standard record templates and associated images to be used in the articles were also collected beforehand. The support of the Wikipedia community was also crucial to the training success. The training sessions and their purpose were announced on the project page on Wikipedia in order to seek the help and support of the Wikipedia community (‘Wikipedia talk: GLAM/State Library of New South Wales/November 2012’ 2012).
SLNSW maintained an ongoing relationship with the Wikipedia community through the development of a project page on Wikipedia (‘Wikipedia:GLAM/State Library of New South Wales’ n.d.). The project page outlined the nature of the project strands as well as providing the account names of project participants. The purpose of the page was in part to provide information to the Wikipedia community on what SLNSW was doing and encourage them to support project staff as they developed their skills with Wikipedia. The page also served as a forum through which the community could engage with SLNSW and participate in edit-a-thons set up by SLNSW to create and edit records on a specific topic, such as World War I (‘World War I edit-a-thons/Australia’ n.d.).

As part of this partnership, a Wikipedian in Residence program was also conducted by SLNSW. This was the first such program to be run by a cultural institution in Australia, in which an experienced Wikipedian was resident at the Library for 14 weeks to support the project. The residency was funded by SLNSW’s eRecords Project and a position description was prepared detailing the Wikipedian’s role. One of the main roles of the Wikipedian was to provide individual training and guidance for staff. Project staff benefited greatly from this and were able to further develop experience in record editing and gain a better understanding of the Wikipedia environment, including the interactions between editors and how articles are interrelated (Rubicon49bce and Aliaretiree 2013). Other tasks included in the residency were to assist in documenting project processes, the upgrading and maintenance of SLNSW’s project page on Wikipedia and consulting on suitable content for contribution to Wikipedia.

Partnership with the public library network has also been an integral aspect of the project. One of the main aims of the project was to engage with the NSW public library network and encourage their use of Wikipedia. It was envisioned that public libraries would be able to use their connections with local studies groups and their expertise with local issues to enrich these records over time, using information gathered from their own archives and research. SLNSW’s outreach to the NSW public library network took various forms. Regional training sessions for public library staff on general Wikipedia usage were conducted by SLNSW. Emails were also regularly sent out to the network detailing the progress of the project as well as alerting the libraries to the digitisation of new titles. Finally, project staff liaised directly with public library staff who wished to create or edit their own Wikipedia records.

A number of records have been enhanced and built on by NSW public library staff as a result of this engagement. The following examples illustrate how SLNSW was able to collaborate with the public library network and use its experience with Wikipedia to assist public library staff in adding content. Prior to the announcement of Trove’s 10-millionth digitised page, a record was especially created for The Leader by SLNSW. This start class article has since been substantially upgraded by the staff at Orange Library, who have added a number of images and a significant amount of additional historical information.

Another two examples demonstrate how the project enabled colleagues to add content where they did not have access to Wikipedia training. The record for Werriwa Times and Goulburn District News was created by SLNSW specifically for a local studies librarian to add content to. The librarian had not yet received Wikipedia training and was
encountering difficulties in creating an article, but found the process of editing an existing article much simpler. A similar example occurred with the *Irrigation Record*. An article for this title was originally created by a public library, but the article was rejected as part of the approvals process for record creation. After discussion with SLNSW, the article content was copied by project staff, who edited it for encyclopaedic style and inserted references as required. References had originally been placed in the record, but not in the preferred format. Once the article was resubmitted, it was approved and published, with credit given to the original author in the edit history. These examples both illustrate the importance of the training undergone by SLNSW staff with Wikimedia Australia.

Wikipedia lists were also created in order to assist public libraries in working on the newspaper records. Once the project was underway, a decision was made to create a comprehensive list on Wikipedia of newspapers that had been published in NSW. This was to serve as a central spot to link into the multitude of records being created. What was not envisioned was how large this list would be. This list of over 2,000 titles now provides an entry for every newspaper known to be published in NSW, along with publication dates, localities and whether the title is still published. A similar list was also prepared for LOTE newspapers.

**Project Outcomes and Continuation**

The work package for the project set benchmarks for how many records were to be created or edited. No benchmarks were set for the number of catalogue links to be created in Wikipedia or for the expected increase in website traffic however. This project was intended to explore and experiment with new ways of doing things and the outcomes would determine benchmarks for the future.

The project created or edited more records than planned for. The project aimed to create or edit records for 79 identified newspapers and 50 collection treasures over 12 weeks. A total of 86 newspaper titles had records either created or edited as part of the Wikipedia Project. As of the 4th September 2013, another 54 newspapers have also been included in Wikipedia as part of a continuation of the work. In addition, the project strand focusing on collection treasures created or edited 105 articles relating to collection treasures.

The creation of Wikipedia content has resulted in a significant increase in the number of links from Wikipedia back to SLNSW’s catalogues and to Trove. As at 4th April 2013, Wikipedia contained 887 links back to SLNSW’s catalogues. By the end of the project on 14th June 2013, there were 1,174, an increase of 32% in two months. This is expected to increase over time. As at 2nd September 2013, Wikipedia contained 10,292 links to digitised newspaper articles and 260 links to individual digitised newspapers on Trove. Unfortunately, it is not known how many links to Trove were included in Wikipedia before the project began.

The statistics gained from the SLNSW’s project show that contributing content to Wikipedia and Wikimedia Commons has resulted in increased traffic to the SLNSW website. From July 2012 to December 2012, the number of visits from Wikipedia to the SLNSW website was 6,463, an average of approximately 1,110 visits per month, and there were only 77 visits in total from Wikimedia sites (an average of 12.8 per month).
From January 2013 to June 2013, the number of visits from Wikipedia increased to 8,472 in total (an increase of over 30%), a monthly average of 1,412 visits. There were 190 visits from Wikimedia sites (including Wikimedia Commons and Wikimedia Outreach), for a monthly average of 31.7 visits. According to Google Analytics figures from NLA, the number of visits from Wikipedia to Trove Digitised Newspaper in August 2012 was 9,367. By August 2013, this was 11,389, an increase of 21%.

This shows a direct correlation between contributing SLNSW content to Wikipedia and to Wikimedia Commons, and increased traffic from these sources to the SLNSW website. Visits from Wikipedia increased sharply in May, following the press release on the project, and in June, when the *1813 Crossing of the Blue Mountains* article was featured on the main page of Wikipedia.

The following figures also indicate that the content contributed by SLNSW is being viewed and edited, as well as leading clients back to SLNSW’s online space:

- As at 4th September 2013, 39 newspaper records had been edited by other Wikipedians.
- As at June 2013, 47 of the 100 images uploaded to Wikimedia Commons had been used 82 times on 64 articles.
- As at June 2013, articles created or edited by project staff had been viewed over 50,000 times.
- The record for the *1813 Crossing of the Blue Mountains* was promoted to “Did You Know” status on the Wikipedia front page and received 8,798 hits in the first 8 hours on the day it was listed.

The Wikipedia Project was intended to be a pilot for developing skills amongst staff and new models of service delivery. As such, an aim of the project was to make the work a part of the regular work of SLNSW. For the course of the project, the focus of staff was to learn how to best work on Wikipedia, including the manner in which records could be created and the nature of the Wikipedian community. The project staff therefore largely worked together as a group at set times. This approach allowed the group to swap ideas and brainstorm solutions to problems as they arose. However, with the completion of the initial project, this work method has ceased. Project staff have reported that they are now confident in their ability to create Wikipedia records. Work on Wikipedia is now completed in their own time in accordance with existing work priorities. In this way, the shift to making the work part of business as normal has been an organic one. Records needing to be created or edited are allocated by a work supervisor, and these are completed by staff when they are able to do so. Over twenty staff are now trained in using Wikipedia as a result of the project. It is envisioned that these staff will gradually train all librarians in contributing to Wikipedia.

**Conclusion**

The Wikipedia Project has been a success for SLNSW. The work package’s deliverables were all met within the project’s 12-week timeframe and the benchmarks set for the number of records to be created or edited were exceeded. The project has developed new skills amongst staff in working with the Wikipedia community and these skills are now
being disseminated to others within the Library. The lessons learnt are also being used to develop procedures and standard workflows for creating content within Wikipedia and establish benchmarks for this work.

The project has also been successful in that SLNSW has established a relationship with Wikimedia Australia with the first Wikipedian in Residence program to be run in Australia. Engagement with public libraries has also been successful with a number of libraries having sought SLNSW’s assistance in creating and editing Wikipedia content.

The figures for website traffic gained from the Wikipedia Project show that contributing content to Wikipedia and placing links within its records can increase the number of visitors to a library website. This project has seen an increase in client traffic coming from Wikipedia to both of SLNSW’s main catalogues and to Trove, and has increased the visibility and discoverability of digitised content. The increase in traffic from Wikipedia to SLNSW library systems has not been as large as that which academic libraries have reported. However, this work has been conducted as a pilot only and was intended as a way of experimenting with new methods of working. Website traffic is expected to increase further, as SLNSW expands its efforts to engage with clients online.

The ongoing creation of Wikipedia records is in large part connected to SLNSW’s current Digital Excellence Program through which SLNSW is digitising its vast collection of treasures, including its newspapers collections. It is planned that as this progresses Wikipedia records will be routinely created or upgraded with links back to the digitised works. This will continue the work of the Innovation Project in enriching the content of Wikipedia and taking the unique collections of SLNSW out to the wider world.
References


