



ikipedia & Museums: Community Curation¹

Australia
Liam Wyatt



*Imagine a world in which **everyone** has free access
to the sum of human knowledge in **their** own language.*

This is the vision upon which Wikipedia is built, a born-digital project with a spirit straight from the European Age of Enlightenment. It was an age when there was a desire to see the world and understand everything in it; an age when collectors and classifiers built the great encyclopedias, dictionaries and museums that still play a central role in our cultures. It was also an age of polymath amateurs, working to shorten the shadow of ignorance just a little bit further. In the



James Murray,
first editor of the
Oxford English Dictionary
in the "Scriptorium" built to
compile the first edition.

Murray represents both
the exclusivist and flattening
approaches to curation as he
received suggestions from
the public but also held
absolute editorial authority.

Picture by
an unknown photographer.

present day those amateurs are very much still at work and many go by the name of Wikipedians: millions of Wikipedians, aggregating our cultural heritage piece by piece into a coherent but ever-changing text. However, the tension between knowledge professionals and interested amateurs remains. That tension and the unnecessary opposition of two groups with passion and conviction - one with the authority and one without - can be seen in debates over the term "curation" and debates about what, if anything, professional curators and Wikipedians have in common. This essay presents three points of curatorial policy that differentiate Wikipedia from professional curation and argues that Wikipedia can be regarded as a meeting ground between the two groups - a place of "community curation".

In recent years "curation" has become a word with contested meaning. At one extreme is an exclusivist understanding that sees the curator as a focal point of cultural understanding, commanding deference as the museum-sensei. At the other extreme is a radical flattening of the term's connotations to the point of being merely synonymous with "selecting". This latter perspective is one that simplifies the art of building, conserving, researching and narrating a collection down to making a playlist. However, neither of these extremes is particularly helpful in envisaging successful museum projects which require community engagement - which museums are increasingly being asked to do.

How one feels about Wikipedia is probably indicative of where an individual stands on the issue of defining the term "curator". The former (exclusivist) group might likely see Wikipedia as a well-intentioned but ultimately unhelpful intrusion into the museum space - akin to the proverbial large enthusiastic dog in the small cluttered room: every time it wags its tail it knocks over a piece of furniture. The latter (flattening) group might likely see Wikipedia as a playpen of the technorati with editorial rules and content guidelines designed to make participation by the interested amateur increasingly difficult.

Between these extreme viewpoints, perhaps unsurprisingly, lies a productive, mutually beneficial relationship. Such a relationship is built upon a collaborative understanding of the idea of curation. It respects expertise but demands engagement; it is focused neither on giving nor taking knowledge but on building a shared understanding.

Wikipedia is frequently described as a product of User-Generated Content (UGC) sitting alongside blogging, social-networking and video sharing websites. However, it is far better understood as a place of Community Curated Works (CCW)². For those in the cultural sector, especially for the professional curator, this differentiation might explain why Wikipedia should be approached differently from other online outreach ventures that your organisation might be involved in. "Community" rather than "user" recognises that Wikipedia is more than merely a series of individuals, it is a movement. The individual Wikipedian is not merely a "user" of a corporation's infrastructure but also potentially the author, reader, reviewer and maintainer of every aspect of the project content, code and community. "Curated" rather than "generated" emphasises that Wikipedians not only add new content but also delete and merge, poke and prod content to build a *better* encyclopedia and not merely a *larger* one. "Works" rather than "content"



signifies that content on Wikipedia is not just a collection of discrete pages but a coherent whole. No one page should stay orphaned from or compete with the rest (either through links, editorial style or content) but should increase the depth and breadth of the larger work.

If Wikipedia were a museum it would be part catalogue, part exhibition, all community curation. Moreover, it has a global breadth and depth of content and an audience of massive scale. That these stars should align for a free, non-profit, knowledge-sharing project makes Wikipedia a perfect place for curators to look when investigating ways to engage with "the people formerly known as the audience"³.

That said, it is not altogether self-evident how to engage with Wikipedia even with the best will in the world. Like any collaborative product the documentation is not consistent (let alone complete!) and the rules are not fixed. Further, like any open community, there is no application form to join. For any risk-averse cultural institution this is daunting but it also opens up many opportunities for those willing to invest the time in building a relationship with the Wikipedia community. There may be no neat listing of what you can do, but equally there is no fixed list of what you cannot do. It is a negotiated relationship either way and the possibilities are quite broad and often unexpected.

Well designed GLAM-Wikimedia⁴ collaborations in the past have gone further than achieving their stated aims. In some cases, they have led to other, largely unanticipated, positive outcomes. In 2008 the *German Bundesarchiv* decided after long negotiations to release copyright in the medium resolution digitised version of thousands of images and to upload those to Wikimedia Commons with attribution back to the originals in their online catalogue⁵. The project's explicit purpose was to increase use and awareness of their little-known photographic collection (through incorporation into relevant Wikipedia articles) and to assist in the task of categorising the subject matter of the images - a dauntingly large and hitherto incomplete task. This project was well designed as it married the advantages of the archive (the collection, its associated records and the expertise) with the advantages of Wikipedia (the encyclopedia, its breadth and depth and the community). Within the year the stated purpose had been largely completed with great success. For example the *Bundesarchiv's* portraits of political leaders had become the headline images for their respective biographies across dozens of Wikipedia language editions and the collection as a whole had been categorised into extremely fine details such as "Category: Public Relations in Forestry in the GDR" and "Category: Black and White Photographic Portraits of Politicians" - categories that are easily imported back to the original collection record. Achieving these intended outcomes was noteworthy enough, but the unintended ones turned out to be even more remarkable ...

By 2011 the images have been added to 65 thousand different articles in over 200 language editions with combined pageviews of more than 100 million per month⁶. Not only were these the expected articles (such as the aforementioned biographies) but *Bundesarchiv* images had begun to be used as illustrations for such disparate subjects as "Random Access Memory", "Magnesium" and even "Leprosy"! Photographs such as these illustrate subjects, not merely objects, giving

Bruce Springsteen performing in Bonn, 1988.

Picture by Thomas Uhlemann.

Deutsches Bundesarchiv (German Federal Archive),
Bild 183-1988-0719-38

One of the images donated by the Bundesarchiv.

Now seen thousands of times
a day including in the article "Rock music".



them a far wider audience and varied contextualisation.

The *Bundesarchiv* has had its images used more widely due to their being the first to provide access. Multimedia (particularly photographs) used in Wikipedia articles are often make-do solutions, chosen from the limited range of options available under an appropriate copyright license to illustrate the given subject. Good quality historical multimedia is especially scarce, so any professional or historical multimedia shared by a cultural institution is very likely to quickly become the default illustration on the encyclopedia. Multimedia offered subsequently from other institutions would need not only to prove its relevance to the article but also that it was superior to the existing image in illustrating the subject at hand. Thus, there is a significant first-mover advantage in sharing good quality historical multimedia.

Due to this marked increase in collection visibility, the clickthrough statistics to the *Bundesarchiv* website instantly jumped, producing a corresponding spike in high-resolution image sales. Moreover, this was no short-lived trend. Statistics for Wikipedia pageviews, *Bundesarchiv* clickthroughs and high-resolution image sales have continued to grow ever since putting paid to the oft-cited fear, "if people can see it over there, they won't bother visiting the institution's website". Instead, the increased visibility "out there" led to a marked increase in the relevance, visitation and sales revenue of the *Bundesarchiv's* own site. Even

less expected, but more valuable, were the copious suggestions for improving metadata - the high number matched by the gloriously pedantic detail of some - a cataloguer's dream⁷. Of course, this meant that the previous problem of too little interest in the collection was replaced by the opposite (but welcome) problem of too much interest.



A plantation in Suriname.
 Watercolor. Painted by G.P.H. Zimmermann.
 One of the images used in a 2010 Tropenmuseum exhibition "Art of Survival: Maroon culture of Suriname"
 that was first digitally restored and provided with multilingual captions on Wikimedia commons.



Even within the relatively tight framework of GLAM-Wikimedia photographic collaboration there are many ways of modifying the project to elicit specific outcomes. The *Tropenmuseum* of Amsterdam has embarked on a project over the last few years that leverages the uniquely global nature of Wikipedia. By sharing its collection of images of early 20th Century Javanese culture, the Indonesian Wikipedia community has set about identifying locations and identifying information that was not known to the curators as well as translating the museum's captions⁸. Some Wikipedia-generated captions have even made it to the museum walls during recent exhibitions. Such a collaboration is both an effective route to lowering the cost of time-consuming work as well as a means of engaging the society from which the collection items originated. In an institution with a collection that spans cultures, these outcomes can be achieved in ways that are neither patronising nor exploitative.

For Wikipedia, the desired outcome of any project is always the same - to increase the scope, reach and quality of "free knowledge". Wikipedia itself is a staunchly non-commercial project so it would seem incongruous that collaborative projects often founder before they begin because of concerns over copyright and the fact that Wikipedia cannot accept any content which restricts commercial re-use. The reason Wikipedia appears to be so pernickety about copyright is that it tries to adhere to the principle that anyone can view the content for no cost and also that they may do whatever they want with the contents. This dual freedom - known as "Gratis & Libre" - is fundamental to the "free knowledge" mission of the community and it is why no content that is "for Wikipedia only", "non-commercial" or "non-derivative" may be used. The only restrictions placed upon any re-use of Wikipedia's texts is that it be attributed and that any changes be shared under the same terms. This very essay includes sections of text incorporated from the policy page

"Wikipedia: Advice for the Cultural Sector" and the whole of it (as it legally must be) is provided under the Creative Commons Attribution Share-Alike license (CC-BY-SA)⁹. Therefore you are free to use this very essay in any way you chose so long as you abide by those two criteria. Most importantly, the Share-Alike clause ensures that what is shared freely once stays free forever and that any other works into which this free-content is integrated become free themselves - hence its nickname as "the viral clause".

There are a few other key policies that differentiate Wikipedia from normal curatorial practice and they bear noting. These are the "three N's" - notability, neutrality and no original research¹⁰. Unlike the aforementioned stance on copyright which is key to the mission of sharing free knowledge, these three policies are instrumentalist ones borne out of necessity. They enable the theory of a mass-collaboration encyclopedia to turn into reality. It is through these policies that Wikipedia addresses the most common criticisms against it - that it is "anti-expert" or "anyone could just write anything".

Wikipedia is a subject-centric encyclopedia: cultural institutions are object-centric. Their respective criteria for growth in their collections are *notability* and *significance*. That is, in Wikipedian usage the word "notable" is a term of art meaning that the subject at hand warrants an article in its own right. It is determined by the existence of reliable third-party sources to verify any statements of fact. "Notability" should not be confused with the term "significance" as used by professional curators for whom it expresses the preservation value of an object. However, even though there could be a "statement of significance" for every item in a collection, this does not mean that Wikipedia should have an article about each and every item. For example, an ancient coin is an *object* and a museum must make an assessment to determine its significance and worthiness for accessioning. However, for Wikipedia, even if the coin is a good representation of its type, it may not in itself be "notable" unless there is significant coverage from reliable sources about that specific ancient coin. On the other hand, the *subject* of that coin's denomination would be notable, and Wikipedia will have an article about that, whilst the museum would not.

Highly significant headline objects in any major cultural collection are also likely to be notable (and therefore worthy of their own Wikipedia article) but there is no direct correlation between the two measures. Indeed, there are copious Wikipedia articles about subjects which would not be considered to have any wider significance at all. Examples include the myriad articles about sporting statistics and individual asteroids or proteins. Whilst it may seem perverse that these topics are considered worthy of stand-alone articles yet many significant objects accessioned into cultural institutions are not, this is merely the consequence of the fact that there are detailed, structured and easily discoverable publications made for each individual subject by recognised experts in the respective fields. Just as notability is not synonymous with significance, equally, non-notability is not synonymous with insignificance. Information about a non-notable subject could perhaps be included as a section of a higher-order article in the encyclopedia.

"Neutrality" is the second key editorial principle that differentiates Wikipedia from the cultural sector. Whilst the encyclopedia attempts in all ways possible to maintain neutrality, by virtue of its position as an official arbiter of cultural heritage, the cultural sector is obliged to make



judgements and tell narratives on behalf of society at large - a process that is inherently non-neutral. Yet, it is this specific divergence that forms the greatest bond of dependency between Wikipedia and the cultural sector. Wikipedia is dependent upon verifiable reliable sources for its information and it is the publication of that information by cultural organisations (most especially the scholarly research about that information) that grounds Wikipedia in reality.

"If anyone can edit, what's stopping people taking my knowledge and just replacing it with their own opinions?" is a very common concern raised by experts and is it precisely the one that the editorial principle of neutrality is designed to address. In practice, the opposite problem is more common - what happens when the editing public make the expert's information *better*? In Wikipedia usage the word neutral is shorthand for "neutral point of view" or "NPOV". The neutral point of view neither sympathises nor disparages its subject; endorses nor opposes specific perspectives. It is not a lack of viewpoint, but is rather a specific, editorially neutral, point of view. Professional curators have gone to much effort in the last few decades to make descriptions and publications engaging and embracing the controversies surrounding an object. This is in contradistinction to past practice of writing simple, uncontroversial (even bland) item descriptions. By insisting on "neutrality" Wikipedia is not asking for enforced simplicity or an avoidance of controversial topics but is rather asking for opposing views (and the editors that hold them) to co-operate to produce a single, as-close-as-we-can-get-to-neutral understanding of a subject that uses reliable sources to verify all claims. For example, Wikipedia's article on the theory of evolution includes a well referenced section on social and cultural responses and therefore also discusses the theory of creationism. The two ideas are not separate articles written by opposing viewers but integrated topics written by encyclopedists.

The final point that differentiates Wikipedia most directly from the cultural sector is the editorial policy of no original research (also known as "NOR"). As primary sites of cultural inquiry, cultural institutions are accustomed to research being undertaken using their collections. Indeed, the level of research output is a measure of their success. Research not only sheds new light on hitherto unloved sections of a collection but is crucial for driving social debate. This process is therefore one of necessary elitism as it requires acquired, analytical skill, scholarship, expertise and earned respect to draw conclusions that are insightful and reliable. However in Wikipedia, where any one person can modify any other's work, the opposite paradigm holds. The real-world credentials of the author are deliberately ignored and the focus is purely on the quality of the output. Of course, the model of blind peer review is common in research as a way of proving the quality of a work independent of the author, but that system still relies on a structure of experts and reputation-driven publications. This is not to suggest that this is a bad thing, indeed it is an extremely good thing, but as "the encyclopedia that anyone can edit" Wikipedia cannot make use of that model. Instead, Wikipedia enforces the dual principles of "Verifiability" and "no original research" in effect to outsource truth¹¹.

All facts included in Wikipedia must be verified to a reliable third party source - the more controversial the statement the more sources are required. Naturally, the definition of "reliable" is highly contestable but this is a contestation that occurs in every professional discipline already and is not unique to Wikipedia. The principal advantage of this "outsourcing" system is that it

obviates all discussions about which editor is more correct than the other because the criterion for including facts is always verifiability - not truth. Experts dipping their toe into the waters of Wikipedia often feel affronted when their offer to write their latest research directly in the encyclopedia is rebuffed with cries of "show us your footnotes". Yet it is clear that without NOR any theory is just as valid as any other and the viability of the collaborative project would degenerate rapidly into bickering of whose version of a truth can garner more popular votes. However with the NOR guideline the published research output from cultural institutions becomes absolutely crucial for Wikipedia as a source of reliable references. Far from competing with experts and research publications Wikipedia (after all, a *tertiary* source) is dependent upon the authoritative voice of cultural institutions to provide citations. Wikipedia's use of this information in turn drives traffic back to the original materials and institutions from whence it came. This cycle draws more people into the deep collections records of cultural institutions than ever before.

Wikipedia is like nothing that has ever gone before it and yet it is a direct successor to the enlightenment tradition of naming, classifying and describing the universe. This essay has itself tried to shed some light on some of those historical correspondences and policy parallels. Whilst very complex, often confusing and constantly changing, it is one of the most important cultural resources ever developed. Wikipedia and the cultural sector are often doing the same thing, for the same reason, in the same medium and for the same audience. Why not do it together?

For more information on how cultural organisations can work with Wikipedia, visit www.glamwiki.org

Hoxne_Hoard.

Some of the Roman spoons from the "Hoxne Hoard" held at the British Museum being shown to Wikipedians during a behind-the-scenes tour. Whilst the individual spoons are not notable, the article about the Hoard as a collection is now arguably the best in the entire encyclopedia. Picture by Fae



Notes

1. CC-BY-SA <http://creativecommons.org/licenses/by-sa/3.0/>
2. This differentiation was first elaborated by former Wikimedia Australia president Brianna Laughter in 2008 <http://brianna.modernthings.org/article/137/community-curated-works-ccv>
3. A phrase articulated by Jay Rosen in 2006 http://journalism.nyu.edu/pubzone/weblogs/pressthink/2006/06/27/pp1_frmr.html
4. GLAM is the acronym used by Wikipedians to describe the professional cultural sector - Galleries, Libraries, Archives and Museums (and sometimes covering related fields of broadcasting or education). Wikimedia, as opposed to Wikipedia, refers to the gamut of sister-projects (and their respective communities) alongside Wikipedia including Wikimedia Commons, WikiBooks, Wiktionary...
5. More information about this project is at <http://commons.wikimedia.org/wiki/Commons:Bundesarchiv>
6. Whilst this data is publicly available, it is quite obscure to find. http://toolserv.org/~magnus/glamorous.php?doit=1&category=Images+from+the+German+Federal+Archive&use_globalusage=1&ns0=1&depth=9 and <http://toolserv.org/~magnus/baglama.php?group=Images+from+the+German+Federal+Archive&date=201105>
7. The “Error reports” page can be seen at http://commons.wikimedia.org/wiki/Commons:Bundesarchiv/Error_reports. Most commentary is in German.
8. The homepage of this project is <http://commons.wikimedia.org/wiki/Commons:Tropenmuseum>
9. The advice page can be found at <http://en.wikipedia.org/wiki/Wikipedia:GLAM> and the full legal information about this copyright license can be found at <http://creativecommons.org/licenses/by-sa/3.0>
10. The full policies for these can be read at www.en.wikipedia.org/wiki/Wikipedia:N, www.en.wikipedia.org/wiki/Wikipedia:NPOV and <http://en.wikipedia.org/wiki/Wikipedia:NOR>
11. This policy can be seen at <http://en.wikipedia.org/wiki/Wikipedia:Verifiability>

The 7th conference of Wikimania 2011, Haifa, Israel. Attendees from all over the world passionate about working together to share knowledge freely. Picture by Wikimedia Israel's photostream

