

インターネット百科事典「ウィキペディア」におけるデジタル情報の利活用方法

荒井翔平（東京薬科大学）

How to Utilize Digital Information in *Wikipedia*

Arai Shohei (Tokyo University of Pharmacy and Life Sciences)

- ・公共財／Public goods
- ・オープンデータ／Open data
- ・クリエイティブ・コモンズ・ライセンス／Creative Commons license
- ・デジタルアーカイブ／Digital archives

はじめに

フリーのオンライン百科事典である「ウィキペディア」は、多くのインターネットユーザーに認知されており、その多様な収録項目からナレッジデータベースで同様の存在価値を示すサイトは皆無と言えるだろう。同時にその情報の存在感が増した分、ウィキペディアに限らずインターネット上に流通されている情報は、価値どころか、存在しないかのように見られてしまうという現状も垣間見える。本稿では、文化財を含めたデジタル情報を利活用するという観点から、ウィキペディアが持っている役割、さらにはそこからどのような情報流通を図ることができるかについて考えたい。

ウィキペディアとウィキメディア財団

ウィキペディア（Wikipedia）は、2001年に開始されたプロジェクトで、その源流はウィキペディアの創始者であるジミー・ウェールズ氏が、サーチエンジンの会社を経営していた1990年代後半に「オンライン百科事典」の実験を実施したことがきっかけである¹⁾。ウィキペディアという名称は、「ウェブブラウザを用いて不特定多数のユーザーが直接コンテンツを編集するウェブサイトを意味する「ウィキ（wiki）」と、百科事典を意味する英単語である“encyclopedia”から成る造語である。前者のウィキシステムはMediaWikiというシステムを用いて稼

働している。MediaWikiについては、後述するが可能な限り簡単に編集できることを見据えて設計された。英語版やその他の言語版ウィキペディアが開始された後、ローマ字（英字）のみを表示する日本語版が設置、その後日本語の文字に対応するようになつた。プロジェクトの開始から20年という節目を迎えた2021年には、日本語版ウィキペディアで130万項目を収録するほどまで成長している²⁾。

（1）ウェブの中のウィキペディア

全世界のウェブサイトの1か月あたりの訪問者数を比較すると表1のとおりである³⁾。Googleの925億アクセスを筆頭にウィキペディア（英語版・日本語版などすべての言語を含む）のアクセス数は61億アクセスで、世界5位である。ウィキペディアが持つ影響力は、かなり大きいといえるだろう。

表1 世界のウェブサイトにおける月次訪問者数³⁾

Rank	Website	Monthly Visitors
1	Google.com	92.5B
2	Youtube.com	34.6B
3	Facebook.com	25.5B
4	Twitter.com	6.6B
5	Wikipedia.org	6.1B

2022年1月現在、英語版を筆頭にウィキペディアは325言語で運営されている⁴⁾。各言語版は独立した運用を担保しており、日本語版も同様である。なお同時に「日本版」ではなく「日本語版」であることに留意されたい。それは、後述する「中立的な観点」に関連するところだが、国（政府）間であったとし

表2 ウィキペディアの各言語版の状況（純記事100万記事以上・2021年1月8日現在）⁴⁾

順位	言語	純記事数	管理者	登録者数	活動中の登録者
1	英語	6,435,185	1,061	42,839,848	118,074
2	セブアノ語	6,104,534	6	85,823	186
3	スウェーデン語	2,747,071	65	803,834	2,272
4	ドイツ語	2,650,941	189	3,842,000	18,551
5	フランス語	2,387,442	160	4,272,459	18,153
6	オランダ語	2,076,847	37	1,184,318	3,954
7	ロシア語	1,784,434	77	3,097,477	11,191
8	スペイン語	1,743,292	65	6,425,404	13,861
9	イタリア語	1,734,706	119	2,202,929	7,985
10	エジプト・アラビア語	1,522,858	7	170,940	213
11	ポーランド語	1,504,026	105	1,143,538	4,319
12	日本語	1,308,203	40	1,878,562	14,917
13	ベトナム語	1,269,689	20	845,161	2,512
14	ワライ語	1,265,630	3	50,499	95
15	中国語	1,249,013	65	3,173,269	8,218
16	アラビア語	1,152,322	26	2,195,915	6,161
17	ウクライナ語	1,131,144	45	590,581	3,108
18	ポルトガル語	1,080,220	66	2,695,072	8,597

表3 ウィキメディア・プロジェクトの種類⁵⁾

名称	URL	創設年	内容
ウィキペディア	wikipedia.org	2001	百科事典
Wiktionary	wiktionary.org	2002	辞書・シソーラス
Wikibooks	wikibooks.org	2003	教育用テキスト・学習用素材
Wikinews	wikinews.org	2004	ニュースの提供
Wikiquotes	wikiquote.org	2003	引用句集
Wikisources	wikisource.org	2003	著作権フリーな文書の収集・翻訳
Wikiversity	wikiversity.org	2006	教育・研究用素材
Wikivoyage	wikivoyage.org	2006*	旅行ガイド
Wikimediacommons	commons.wikimedia.org	2004	メディアファイルの収集庫
Wikimediacommons	incubator.wikimedia.org	2006	新規言語版プロジェクトの試験・開発
Wikimeta	meta.wikimedia.org	2001	Wikimedia・プロジェクトに関する議論の場
Wikispecies	species.wikimedia.org	2004	生物分類目録
Wikidata	wikidata.org	2012	知識データベース
Wikimania	wikimania.wikimedia.org	2005	Wikimaniaの会議

* 2013年よりWikimedia・プロジェクト

ても中立的な内容収録を行うという方針に基づいている。

純記事数をベースに各言語版を比較すると表2の通りとなる⁴⁾。なお、ここでいう純記事とは、別項目への転送（リダイレクト）記事を含まない、かつWikimediaの別記事へのリンク（内部リンク）を1つ以上含むページを示す。

(2) ウィキメディア財団

Wikimediaを運営しているのは、アメリカの非営利組織であるWikimedia財団（Wikimedia Foundation）である。財団として雇用した従業員が500名ほどいるが、各言語版で活躍する執筆者や管理者は基本的に無償で参加するボランティアである。ここには各言語版で活動している管理者も含まれる。また、財団は企業・個人からの寄付で運営し

表4 主なウィキ記法の種類

記入方法	用途	備考
[[半角鍵カッコ]]	記事内リンク	姉妹プロジェクトへも内部リンクで対応可
[http://url/ title]	外部リンク	
<ref>～</ref>	脚注情報	
== 半角イコール ==	見出し	小見出しごとにイコールの数が増える

ており、その費用をサーバー運営費や活動助成などに充てている。

これは、各プロジェクトが中立的に運営されることを目的にしているためである。つまり「ある企業からスポンサー料を受領した結果、広告主に対して、忖度した記事が作成される可能性」を排除している。ウィキメディア・プロジェクトが広告を掲載せず、寄付の依頼を頻繁に掲載するのはそのためである。

(3) ウィキメディア・プロジェクト

ウィキメディア財団は、百科事典プロジェクトであるウィキペディア以外にも、教科書を作るウィキブックスや、画像や音声・動画を集積するウィキメディア・コモンズなど、多くのプロジェクトを運営・展開している⁵⁾。実際にはそれぞれのプロジェクト間は連携しており、相互に利活用できるようにソフトウェアが設計されている。これらのプロジェクトを総称して「ウィキメディア・プロジェクト」と呼称している。プロジェクトの一覧については、表3に示した。相互の利活用例として、ウィキペディアの記事中に使用する画像については、メディアファイルの収集を主目的とする「ウィキメディア・コモンズ」にアップロードすることで、ウィキペディアを含めた世界全言語版のウィキメディア・プロジェクトで利用することが可能となる。

考え方とライセンス

ここではウィキペディアの編集における基本的な考え方について触れておきたい。基本原則として「5本の柱⁶⁾」、記事の内容に関して「三大方針」がある。いずれもウィキペディアを利活用するうえで理解しておかなくてはならない考え方である。また、これ

を反映する著作権条件としてライセンスを設定している。

(1) 5本の柱と三大方針

5本の柱は、「百科事典である」「中立的な観点に基づく」「利用はフリーで、誰でも編集可能」「行動規範がある」「これ以外に確固としたルールはない」というウィキペディアの原則を示している⁶⁾。そのうち、「中立的な観点」を相互補完する形で「検証可能性」「独自研究は載せない」という2つが提言され、これらが三大方針として設定された。ウィキペディアの記事が構成されていくうえで“三大方針は議論の余地がないもの”としている⁷⁾。これらの方針を遵守しながら記事を作成していくために必要なのが、図書館や文書館などに所蔵されている、出典となる資料である。

(2) ライセンス

作成した文章や画像は作成者・撮影者の著作物であるが、前述のとおりウィキペディアでは二次利用を促進する観点から、クリエイティブ・コモンズ・ライセンス (CC-BY-SA) と GFDL のデュアルライセンスを導入している⁸⁾。これは、作成されたコンテンツが誰でも自由に複製・改変・再配布できる枠組みを構築して二次利用を促進している。各人がウィキペディアの記事編集を終え、保存する際には、このライセンス条件に同意することとし、編集者一人一人の許諾を得たうえで公開される。

(3) ウィキシステムでの編集

ウィキペディアをはじめとするウィキプロジェクトで利用しているソフトウェアである Mediawiki は、HTMLなどのプログラム言語を習得していない利用者であっても記事編集に参加しやすくなることを念頭に「ウィキ記法」を用いたシステム（マーク

アップ) が構築されている。主な内容は表 4 に示すとおりだが、これ以外にもスタイル表記のための記法などがある。また、後発ではあるが、ウィキ記法すら理解していなくても編集可能なシステムとして開発・実現したのが「ビジュアルエディター」である。マウスとキーボードの直感で操作できるように設計された。

アウトリーチとしての利活用

GLAM⁹⁾ とウィキメディア・プロジェクトもしくはウィキメディア財団との連携は、世界各国で展開されている。例えば、美術館が持っている古美術、図書館で所蔵されている古い絵巻や古地図を写真で撮影し、それにクリエイティブ・コモンズ・ライセンスを付与し、ウィキメディア・コモンズにアップロードするという取り組みや、公文書館で所蔵されている文書を電子化してウィキソースに保管するなど、それぞれの館での所蔵状況に合わせて、可能なところから取り組んでいる。その例の一つとして、大英博物館のアーカイブから相当数のメディアファイルが、ウィキメディア・コモンズにアップロードされ、ウィキペディアなどのプロジェクトに限らず、全世界で誰もが自由に利用できるようになっている¹⁰⁾。ほかにも、イスラエルの大学図書館とプロジェクトを行う¹¹⁾など、様々な地域で多様なプロジェクトを展開している。国内でも、同様のオープンアクセス化が順次進んでいるものの、その後の利用については、進展しているとは言えない状況である¹²⁾¹³⁾。単純にデジタルアーカイブとして公開するだけではなく、世界中から自由に利活用できるシステムとして GLAM 関係者がウィキメディア・プロジェクトの存在を認知し、ウィキメディア・コモンズにデータをアップロードしてもらう必要性もあるが、その認識はまだ広がってはいない。

つぎに、地域社会との連携については、これまで多くの実践が行われてきた。GLAM に限らず、地域が持つ文化財情報は多岐にわたっているが、デジタル化されているデータはかなり限定的であり、逆

にデジタル化されていないデータは、デジタル化されるタイミングを待ちわびている状況である。しかしながら、自治体予算やそこに定常的に割ける労力などの限界により、デジタル化が進まないということも多いことから、ウィキペディアタウンなどの取り組みと連動させ、デジタル化しようという動きもある。そのほか、教育機関との事例については、授業¹⁴⁾¹⁵⁾や学校図書館との連携¹⁶⁾、大学の講義¹⁷⁾など、その事例は多岐にわたる。今後ウィキペディアとの連携が模索されるのは、教育連携のさらなる深堀りであろう。前述したとおり、方針順守の観点から資料を元に記事を作成していくが、これは、情報リテラシー教育の手段で使われるべきである。ただ、大学の講義と図書館の連携でさえ、それが生かされていないと考えられるため¹⁸⁾、教育の一環、特に学習指導要領との連携として図書館とウィキペディアをどう扱うのかについては、さらなる議論の必要がある。なお、直近における教育とウィキペディアの連携については、佐々木部分を参照されたい。

知識の財産としてのウィキペディア

デジタル資料の利活用の機会は幾倍にも増すだろう。もちろんその他の媒体でも活用は可能だ。また、その資料が研究やメディアによって再構成される可能性も十二分にある。

様々な情報のデジタル化により、利活用の機会は幾倍にも増すだろう。例えば、様々な事情により書籍には掲載されることのなかった資料がウィキメディア・コモンズで公開されることで、その資料を用いたウィキペディア記事の作成が可能になる。もちろんその他の媒体でも活用は可能だ。また、その資料が研究やメディアの再構成などにより、知識の知見を生むことも期待される。これまで、ウィキメディア・プロジェクトがデジタルアーカイブの側面から事例などを紹介してきたが、ウィキペディアにある文面の原点は、すべては研究の積み重ねから成立している。そのベースとなる情報は、書籍などの

紙情報が中心だが、電子情報が検索・認知しやすいところに存在すること、同時に書籍には様々な事情により掲載されることができなかつたデータが電子媒体となって公開されることによって、その利活用の速度は幾倍にも早まると同時に、ウィキペディアに限らず、世界中で利用が可能になるのが、デジタル情報だと考える。研究や取材の成果で出来上がった情報が、ウィキペディアを編集する誰かに渡り、様々な情報を要約すると同時に分類することから、記事の編集が始まる。つまり、ウィキペディアの記事が「知識の財産」として、情報の再生産となるのだが、それまでにも多量の情報から研究が行われているということが、研究者側から意外に認知されていないことも現状ウィキペディアが抱える課題の一つである。同時に、ウィキペディアをはじめとする上記のウィキメディア・プロジェクトが、このような可能性に満ちたメディアであることはあまり認識されていないと思われる。資料の利活用を促進し、情報をまとめたツールとしてウィキメディア・プロジェクトは有用であることを、特に研究者コミュニティに対し積極的にアピールしていく必要があるだろう。

さいごに

ウィキペディアタウンが日本国内では2013年に横浜で初めて開催された以降、様々な取り組みが全国各地で行われている^{19) 20)}。その中で、文化財情報をはじめとした様々なデータがデジタル化されていない状態で埋没している資料や、デジタル化されたとしても公開されずに保管されている資料、公開しているが横断検索にヒットしない資料などの利活用の可能性を、ウィキメディア・プロジェクトを通じて、高められるようにする必要があると考えている。ただ、このような状況が続いてしまうようでは、利活用を待っていた情報は埋没したまま年月が経過してしまう可能性が非常に高い。そのためにもできるだけ早く実際にデータをウェブ上にアップロードされるように働きかけるよう進めていただきたいと

強く願う。

謝辞

本稿は、文化財とウィキペディアがどのように融合していくかという問い合わせ特集構成していただいた、独立行政法人国立文化財機構 奈良文化財研究所の高田祐一氏・野口淳氏の両氏の発案で実現した。同時に、文化財情報の利活用とその発展に関する考え方をまとめていただきたいとのご意見を青木和人氏や長野市立更北中学校の佐々木宏展氏から頂戴し、僭越ながらその視点を加えたものである。また、長野県高森町立高森北小学校の宮澤優子氏、香川短期大学の中俣保志教授には、現地訪問時の対応や各種事例に関するご紹介に限らず、公私を問わず多大なるご支援を賜った。同時に、長野県伊那市地域おこし協力隊の諸田和幸氏、早稲田大学ウィキペディアンサークルのみなさまには本稿執筆にあたり、様々な視点からご意見を頂戴した。

最後に、所属先である東京薬科大学の皆様には、コロナ禍で世間が暗い雰囲気に落ち込んでいる状況であっても、常に明るく勇気づけていただいた。この場を借り、心から感謝申し上げる。

【補註および参考文献】

- 1) 「ウィキペディア初の項目って? 創設者が明かす
草創期: How I built this 日本語版(中)」『NIKKEI
STYLE キャリア キャリアコラム』2021年12月6日、
<https://style.nikkei.com/article/DGXZQOLM095R60Z01C21A1000000/> (2022-01-01 確認)
- 2) 「Wikipedia: 発表/130万項目達成」『ウィキペディア
日本語版』2021年11月26日06:06 (UTC)版、<https://ja.wikipedia.org/w/index.php?oldid=86718014>
- 3) "The 50 Most Visited Websites in the World" 2021-01-27, Visual Capitalist, <https://www.visualcapitalist.com/the-50-most-visited-websites-in-the-world/> (2022-01-01 確認)
- 4) 「Wikipedia: 全言語版の統計」『ウィキペディア
日本語版』2022年1月8日02:55 (UTC)版、<https://ja.wikipedia.org/w/index.php?title=%E6%97%A5%E8%AA%9E%E6%8B%9B%EF%BC%88%E7%84%A1%EF%BC%89&oldid=10910000>

- ja.wikipedia.org/w/index.php?oldid=87408939
- 5) 「Wikipedia：ウィキメディア・プロジェクト」『ウィキペディア日本語版』2021年10月3日14:04 (UTC) 版、
<https://ja.wikipedia.org/w/index.php?oldid=85868688>
- 6) 「Wikipedia：五本の柱」『ウィキペディア日本語版』
2021年2月14日16:41 (UTC) 版、<https://ja.wikipedia.org/w/index.php?oldid=81835740>
- 7) 「Wikipedia：中立的な観点」『ウィキペディア日本語版』
2021年8月9日09:24 (UTC) 版、<https://ja.wikipedia.org/w/index.php?oldid=84955809>
- 8) ウィキペディアのコンテンツを二次利用する場合にはどちらのライセンス条件を用いてもよい。
- 9) GLAMとは、美術館 (Galleries)・図書館 (Libraries)・
公文書館 (Archives)・博物館 (Museums) を集約して現わす。
- 10) “Category:Collections of the British Museum by object type” https://commons.wikimedia.org/wiki/Category:Collections_of_the_British_Museum_by_object_type (2022-01-01 閲覧)
- 11) 国立国会図書館関西館図書館協力課調査情報係
2013 「Wikipedia と GLAM の協力：ハイファ大学図書館の事例（イスラエル）」『カレントアウェアネス-R』<https://current.ndl.go.jp/node/23301> (2022-01-01 閲覧)
- 12) 東修作 2020 「オープンアクセス画像の構造化データベースとしてのウィキメディア・コモンズの活用：WMFによる画像インポートに関わる法的な整理例とデータ整備の試みの紹介」『デジタルアーカイブ学会誌』第4巻第1号 pp.37-40
- 13) 福島幸宏 2020 「文化財情報を真の公共財とするため」『奈良文化財研究所研究報告』第24冊 pp.118-121
- 14) 八田友和 2021 「地域の人的・物的資源を活用した授業の一考察：芦屋市におけるウィキペディアタウン活動を事例に」『関西教職教育研究』第10号 pp.55-61
- 15) 市川博之 2017 「Wikipedia Town in 下田小学校&白浜小学校」『先生のための授業に役立つ学校図書館活動データベース』東京学芸大学学校図書館運営専門委員会 <http://www.u-gakugei.ac.jp/~schoolib/htdocs/index.php?key=jos82vv6k-121> (2022-01-01 閲覧)
- 16) 伊達深雪 2019 「学校図書館による「地域探究」授業および活動の支援：地域と共に開催するウィキペディアタウンより」『図書館雑誌』第113巻第12号 pp.804-805
- 17) 日向良和 2021 「ウィキペディアタウンを通じた地域情報の発信：博物館情報メディア論の一環として」『博物館研究』第56巻第1号 pp.9-13
- 18) 荒井翔平・堀口健・林真紀 2013 「図書館を学生と共に創るための活動—モニター活動の取り組みから見えること」『東京都市大学横浜キャンパス情報メディアジャーナル』第14号 pp.97-101
- 19) 「「ウィキペディア」の編集を通して地域を知り愛着育む 池田町でワークショップ」『北アルプス経済新聞』2021年11月29日、<https://kita-alps.keizai.biz/headline/174/> (2022-01-01 閲覧)
- 20) 青木和人 2022 「ウィキペディアタウンからウィキペディア文化財へ」『奈良文化財研究所研究報告』第29冊 pp.51-56