

Wikipediaにおける情報の質

鈴木 優^{1,a)}

受付日 2013年3月20日, 採録日 2013年6月28日

概要: 本論文は, Wikipedia に記述されている情報の質に関する研究のサーベイである. Wikipedia は, 不特定多数の利用者によって容易に編集を行うことができるという特徴から, 従来の紙媒体による百科事典と比較して速報性が高いことや, 広範囲な情報が得られる利点がある一方で, 質が低い情報が含まれるという欠点も存在する. 利用者は, 必ずしもすべての情報に対して質が高い情報かどうかを適切に判定することができないため, Wikipedia に対して質の高さを測定しようという試みが数多く行われている. 本論文では, 質の定義についての議論を行ったうえで, 現在までに行われている Wikipedia に関する調査について紹介し, Wikipedia に含まれる情報に対して, 情報の質を測定する研究について紹介する.

キーワード: CGM, UGC, Wikipedia, 質, 信憑性, 貢献度, 相互評価

Measuring Quality of Wikipedia Articles: A Survey

YU SUZUKI^{1,a)}

Received: March 20, 2013, Accepted: June 28, 2013

Abstract: Wikipedia is one of the most successful and well-known User Generated Content (UGC) websites. Because any user can edit any article, Wikipedia has more and fresher information than existing paper-based encyclopedias. Many experts submit texts to Wikipedia, and the texts should be informative for readers. However, these texts are not reviewed by experts, then the number of poor quality texts are also dramatically increase. On the other hand, many readers cannot easily identify texts which are good quality or not, because not all readers are experts. In this paper, we survey the studies for assessing a quality of Wikipedia articles.

Keywords: CGM, UGC, Wikipedia, quality, credibility, trustworthiness, expertise, peer review

1. はじめに

Wikipedia^{*1}はインターネット上の百科事典であり, 集合知を用いたサイトの中でも最も成功したサイトの1つである. Wikipedia は一般の百科事典とは異なり, 利用者であれば誰でも記事を編集することができる UGC (User Generated Content; 利用者生成コンテンツ) の1つである. 2001年1月に英語版 Wikipedia が運用を開始し, 同年5月に日本語版 Wikipedia が運用を開始した. その後, 編集者の数や記事の数は指数関数的に増加し [1], 2013年5月には英語版 Wikipedia に約 422 万件, 日本語版に約 86 万件の記事^{*2}が存在している^{*3}. このように, Wikipedia

の規模は日々拡大しており, インターネット上の重要な知識源となりつつある.

Wikipedia に含まれるデータは, 情報収集目的以外に研究目的としても数多く利用されている. これは, 編集履歴やタイトル一覧, 編集者一覧^{*4}, 閲覧履歴^{*5}などの多様な情報が公開されており, 誰でも利用が可能であることが理由の1つとしてあげられる. Wikipedia に関する論文は, 2002年に初めて発表された [2] 後, Wikipedia に掲載され

¹ 名古屋大学
Nagoya University, Nagoya, Aichi 464-8601, Japan
^{a)} suzuki@db.ss.is.nagoya-u.ac.jp

^{*1} Wikipedia: <http://www.wikipedia.org/>
^{*2} 標準名前空間にあって, リダイレクトではなく, 少なくとも1つの Wikipedia 内リンクを持つページのこと. <http://ja.wikipedia.org/wiki/Help:記事とは何か>
^{*3} 全言語版の統計: <http://ja.wikipedia.org/wiki/Wikipedia:全言語版の統計>
^{*4} Wikimedia Downloads: <http://dumps.wikimedia.org/>
^{*5} Page view statistics for Wikimedia projects: <http://dumps.wikimedia.org/other/pagecounts-raw/>

ている論文^{*6}だけでも、2010年には年間160件程度が学術雑誌などに掲載されている。Wikipediaのデータを利用した研究は情報視覚化、データマイニング、言語処理などの情報科学分野にとどまらず、社会学[3]やメディアリテラシ[4]など多岐の学術分野にわたっている。Wikipediaを運営しているMediaWiki財団自身も、Wikipediaや関連サイトに関する研究コミュニティ^{*7}を構築していたり、MediaWiki財団が主催するWikimania^{*8}や、Wikiを中心とした国際会議であるWikiSym^{*9}を開催したりするなど、今後ますますWikipediaを含むUGCに関する研究は増加すると考えられる。

Wikipediaに関する研究は、大きく2つに分類することができる。1つはWikipediaを情報源とした新たな知的資源の構築を目的とした研究であり、中山ら[5]によって紹介されている。もう1つはWikipediaそのものの態様を明らかにしようとする研究であり、我々の知る限りこれらの研究についてまとめられたものは存在しない。そこで、本論文では後者の分類に属する研究を扱う。

情報そのものの質を測定する研究において、測定対象としてWeb文書や電子メール、SNS(Social Network Service)など、Wikipedia以外の様々な電子文書を扱った研究が存在する。これらの情報に関する信頼性については松林ら[6]がまとめており、特に大島ら[7]は、Web上の情報に対する信憑性についての考え方、および検証技術に関して述べている。本論文では、質を測定する対象を特にWikipediaを対象としたものに限定することによって、Webページから得られないような多くの情報源が存在したときに、これらの情報源から質を測定する手法について紹介する。

本論文で扱う言語版は特に限定していないが、多くの研究では英語版を対象としている一方で、記事数が多く扱いにくいことから、それぞれの研究者の母国語版を利用している場合もある。Wikipediaは言語版ごとにある程度独立して運営されていること、それぞれの言語版に対して記述している編集者の大部分は異なると考えられること、各言語版に対応する国ではそれぞれ文化が異なる場合が多いことなどからも、ある言語版で成り立つ知見が他の言語版で成り立たない場合が考えられる。ところが、言語版の違いに着目した研究は少ないため、本論文では特に示さない限り、紹介する論文において言語版の差異について考慮していない。

本論文の構成は次のとおりである。まず2章において、情報の質と信憑性の定義について述べる。次に、3章においてWikipediaに関する様々な調査について紹介し、4章においてWikipediaの記事に対する情報の質の測定方法に

ついて紹介する。最後に5章において、Wikipediaに関する研究の現状および今後の展開などについて述べる。

2. 情報の質、信憑性の定義

Wikipediaが成功していると多くの人に考えられている根拠の1つとして、その情報が十分高品質であると多くの人に認識されていることがあげられる。アメリカでは、裁判の際にも参照されたことがある^{*10}ほどであるが、過去の研究において、一般の百科事典と比較して十分高い質があるという意見[8]もあれば、そうではないという意見[9]もある。

情報の質という概念は曖昧であると感じられる概念であり、定義を行うことが難しい。また、質が高いことと真実、事実であることは、直感的には同義であると考えられることがあるが、この直感的な概念は必ずしも正しくない。なぜなら、事実や真実とは何かを厳密に定義することは困難であるため、記述が真実であるかどうかを判定することもまた困難であるためである[10],[11],[12]。そのため、情報の質に関する概念は、測定された質の利用目的やデータそのものの性質などに応じて、適切な定義がなされなければならない。また、計算機上で質を扱うためには、モデル化することができることも重要な要因である。そこでまず、従来の研究において質がどのように定義されているかについて述べる。

まず、質に関する一般的な議論を行う。国語辞典の1つである広辞苑[13]では、質を「対象が他の対象と区別する特色となっているもの」と定義している。つまり、人間は対象をどのように区別しているかは人それぞれであることから、質は人間の主観による性質であり、同一の記述であっても人によって異なる判定を行う場合があることが分かる。そのため、ある利用者が情報の良否をどのように決めるか、信用するかどうかを計算機が自動的に推定することは困難であるため、情報に対して質を直接測定することは難しい。そこで計算機上で情報の質を測定するためには、質を抽象化する必要がある。つまり情報科学において、質を測定する研究とは、本質的には質をモデル化する手法であるといえる。

現在までに、様々な分野で質の定義を行う試みがなされている。心理学では、信憑性にはtrustworthinessとexpertiseの2つの側面があり、それぞれ主観的、客観的な側面があると考えられている[14],[15]。Trustworthinessはbelievabilityと言い換えられることもあり、記述が多くの閲覧者によって信じられているかどうかを示す値であり、主観的な値であることが多い。Expertiseは専門性と言い換えることができ、主観的に認知することはあるが、参照されている情報資源などの性質などから、比較的客観的に認

^{*6} Wikipedia in Academic Studies: http://en.wikipedia.org/wiki/Wikipedia:Wikipedia_in_academic_studies

^{*7} Research: <http://meta.wikimedia.org/wiki/Research:Index>

^{*8} Wikimania: <http://wikimania.wikimedia.org/>

^{*9} WikiSym: <http://www.wikisym.org/>

^{*10} Courts Turn to Wikipedia, but Selectively: <http://www.nytimes.com/2007/01/29/technology/29wikipedia.html>

知することができる場合もある [16]. ところが, Wikipedia の記事に対してこれら 2 つの指標を用いた質の測定することはきわめて困難である. なぜなら, これらの指標による測定において, 編集者自身にどのような知識があるのか, 考え方にどのようなバイアスがあるのかなど, 編集者に関する様々な情報を収集する必要があるが, Wikipedia ではこれらの情報が収集されていないためである. そのため, 4.1 節であげる人手による質の測定に関する研究では, この定義そのものではなく, この定義に準じた質の定義を用いていることが多い.

Wand ら [11] は, 映画の感想など実世界データを対象に質を測定する研究についてのサーベイを行っているが, 多種多様な文書に対する質の定義について, 共通した合意をとることが難しいと述べている. つまり, Wikipedia で扱っている情報のうちの一部の分野であっても, 質の定義は様々であることが分かる.

紙媒体の百科事典においても, 質を測定する試みが行われている. Wong [17] は紙による百科事典の評価や歴史について調査したうえで, 百科事典の外形的な質を評価する際特有の指標として, 次の 9 種類: (1) 取り扱う範囲 (Scope) (目的, 主題, 網羅性, 聴衆, 配置, 種類), (2) 構成 (Format), (3) 独自性 (Uniqueness), (4) 権威 (Authority), (5) 正確性 (Accuracy) (正確性, 信頼性, 客観性), (6) 最新性 (Currency), (7) 使いやすさ (Accessibility) (索引), (8) 利用者の要求との関連度 (Relevance), (9) 価格 (Cost) を提案している. これらの指標の一部は紙媒体である百科事典に特有の指標であるため, オンライン百科事典ならではの指標として重要である, 記事を作成した編集者間の関係や編集履歴などの指標は入れられていない. ただ, 取り扱う範囲や構成, 独自性など一部の指標は, Wikipedia の質を評価するうえで参考となる指標であるといえる.

情報の質を, 情報が真実であるかどうかという観点から検証した研究もある. Sheppard ら [18] は, Wikipedia などのオープンコラボレーションにおける質の定義について述べている. 彼らは市民科学^{*11} コミュニティ構成員に対してアンケートを行い, 信憑性 (Reliability, Trustworthiness), 正確性 (Accuracy), 整合性 (Consistency) が情報の質を決定するための重要な要素であることを示している. この研究では質を真実であるかどうかという観点で検証しており, 情報がない状態と誤った情報が存在する場合に, 情報がない状態のほうが良い状態であると述べている. なぜならば, 誤った情報であるかどうかを確認する作業は手間がかかり, 修正を行うことは容易ではないためである.

最後に, Wikipedia 自身では記事の質をどのように扱っているかについて述べる. Wikipedia では, 完璧な記事^{*12}や

秀逸な記事^{*13}に関する記事において, 高品質な記事の条件が記述されている. ところが実際には, これらの条件の記事が満たしているかどうかを判定することは困難であり, 人によって感覚が異なる. Wikipedia では, 一定の基準を満たした編集者に対して投票権を与え, 記事に対して投票を行うことによって, 秀逸な記事や良質な記事であるかどうかを判定している. これは 4.2.1 項で述べている投票による手法であり, 質の測定精度が高いことを期待できるが, 網羅性に欠けると考えられ, 質が高い記事が必ずしもすべて秀逸な記事や良質な記事であると判定されているわけではない.

以下の章では, 質と信憑性という単語を同じ意味で用いている. この章で述べた議論においては, 確かに質と信憑性は異なる概念であり, 信憑性は質という大きな概念の一部であるという認識が主である. ところが実際には, これらの値を測定する段階では区別されていないことも多い. また, 質と信憑性の差異を意識した研究は非常に少ないのが現状である. 一方, それぞれの測定手法相互には密接な関係があるため, これらの関係を明示することは有用であると考えられる. そこで, 以下の章では特に区別を行わなければならない場合を除いて, どちらも質と表現することとする. また, 多くの研究において trustworthiness のことを質と示すことが多いため, 特に示さない限り, 質は trustworthiness のことを指す.

3. Wikipedia に関する調査

Wikipedia における情報の質を測定するためには, Wikipedia の記事から質に影響すると考えられる特徴を抽出する必要がある. ところが, どのような特徴を抽出することができるかが分からなければ, どの特徴を質の測定に活用することができるかを選択することは困難である. そこで, 従来の研究で行われてきた, Wikipedia に関する調査および調査から得られる特徴について紹介する. まず, 3.1 節において閲覧者の利用状況, 3.2 節において記事の状況について紹介する. そして, 3.3 節において編集者がなぜ Wikipedia に貢献するのかという点について紹介を行い, 最後に, 3.4 節において様々な利用者によって記述されている Wikipedia がどのように統治されているかについて紹介する.

3.1 閲覧者の利用状況

Wikipedia の利用状況に関する調査は数多く行われている. たとえば 1 章の冒頭でも述べたように, Wikipedia を運営する Wikimedia 財団によって, 記事数, 投稿数, 記事の閲覧回数などが公開されている. ところが, これらの統計データだけでは具体的な利用者の状況について観察する

*11 科学者ではない一般市民による科学研究.

*12 完璧な記事: <http://ja.wikipedia.org/wiki/Wikipedia:完璧な記事>

*13 秀逸な記事の選考: <http://ja.wikipedia.org/wiki/Wikipedia:秀逸な記事の選考>

ことが難しい。そこで、閲覧者の利用状況に関する、より具体的な調査が行われている。

White により、2007 年現在一般に公開されている Web 2.0 サービス^{*14}の利用者属性に関する調査 [19] が行われている。この調査によると、Web サービス利用者のうち世代を問わず 70%から 84%程度の利用者が Wikipedia を利用している。これは Blog が 40%から 60%程度、Facebook が 5%から 20%程度の利用者によって利用されていることや、この調査で対象となった Web 2.0 サービスのうち最も利用者数が多いことから、Wikipedia の利用者数は相対的に多いことが分かる。また、利用用途は学習や仕事に 60%以上の割合で用いられており、Blog が学習や仕事に利用される割合が 20%程度であること、ほとんどのサービスが 20%以下であることが報告されている。また、調査対象 Web 2.0 サービスのうち、学生や研究者が参考文献として利用しているサービスは Wikipedia が最も多い。

Wikipedia が一般の利用者にとってどのように受け止められているかについて、調査が行われている。Flanagin ら [20] は、Wikipedia の記事が信じられているかどうかを世代によって異なるかどうかを調査した。彼らは、Wikipedia の利用者を 11 歳から 18 歳までの 183 人の若年層グループと 18 歳以上の 283 人の成人グループという 2 つのグループに分けて、それぞれの利用者に対してアンケートを行った。比較対象としてブリタニカ国際百科事典と *Citizendium*^{*15}を設定している。その結果、若年層グループの利用者は Wikipedia の記事を成人グループと比べて比較的信じていないことが分かっている。また、どちらのグループの利用者も 70%から 80%の割合で Wikipedia が UGC であることを理解しておらず、特に成人グループではその傾向が顕著であったことを報告している。

3.2 記事の状況

Wikipedia の記事についての状況を把握することは、記事の質を測定する際に有用な情報であるといえる。そこで現在までに、記事の編集履歴を利用して状況を測定、可視化する方法について研究されている。

Wikipedia に含まれる記事から得られる統計量は、数多く測定されている。まず、Wikipedia を運営する Wikimedia 財団から、記事数や総項目数などが公開されている。さらに詳細な統計量については Holloway ら [21] や Ortega ら [22]^{*16}によって調査されている。

^{*14} ソーシャルブックマーク、オンラインカレンダー、画像共有、共同編集、映像共有、ソーシャルネットワーキング、Blog、ファイル共有、コミュニケーション、およびソーシャルゲームの 10 種類、39 個の Web サービス。Wikipedia は共同編集に含まれている。

^{*15} Wikipedia と同じ UGC による百科事典であるが、編集者は匿名で記事を記述することができない、編集を行う際には身元を身分証明証などで証明する必要があるなど、編集者に制約を課している。査読制度がある。http://en.citizendium.org/

^{*16} WikiXRay: http://meta.wikimedia.org/wiki/WikiXRay

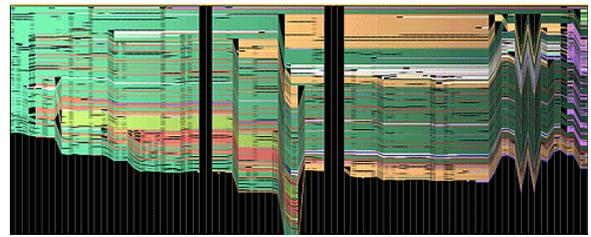


図 1 History Flow^{*19}

Fig. 1 History Flow.

一方、文字数などの簡易な基準を用いた調査も行われている。Kittur らの 2007 年の調査 [23] によると、英語版 Wikipedia に含まれる 147,360 件の記事のうち、78.6%が “B-class”^{*17} (未完成で、さらに信頼できる情報源が必要な状態) 以下の完成度であると報告している。

Viégas ら [24] は、編集者がシステムへの貢献割合や記述量の変化を視覚化するために、編集履歴を History Flow と呼ばれる方法で視覚化した。図 1 に、実際の Wikipedia において編集履歴を視覚化した例を示す。この図は記事「Abortion (妊娠中絶)^{*18}」を視覚化した図である。一番左が最も古いバージョンを表しており、右にいくに従って新しいバージョンを表している。また、各バージョンにおいてどの編集者がどの部分をどの程度記述しているかを色によって表現している。この図では、記事に対して大きな貢献を行った編集者を表す色が、図の中で大きな面積を占めている。そのため、この図からどの編集者がどの程度記事に対して貢献したかを読み取ることができる。また、すべての記事を削除するような問題行為 (Vandalism) が存在するとき、History Flow では黒い割れ目として表現される。そのため、利用者は編集履歴の概観を容易に把握することができる。

Sabel [25] は記事の編集履歴を、木構造によって表現した。編集履歴は通常、記事が編集された時刻によって時系列順に 1 次元で扱われることが多い。ところが、記事は必ずしも単調増加するわけではなく、2 つの異なる版に分岐することも考えられる。Sabel は、記事の版をノードとし、作成された記事内容の類似度を編集距離 [26] によって算出し、類似しているノード間にエッジを構築した。このグラフによって、編集合戦など編集の様子を簡単に視覚化することができる。

Roth ら [27] は、知識ネットワーク解析 (Epistemic Network Analysis) [28] を利用して Wikipedia を解析し、どのような要因が編集者数や記事内容の増加と関係があるかを調査している。その結果、編集者数あたりの編集回数が増

^{*17} Wikipedia: WikiProject Articles for creation/Grading scheme: http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Articles_for_creation/Grading_scheme

^{*18} Abortion: http://en.wikipedia.org/wiki/Abortion

^{*19} 文献 [24] の編集者の Web ページ (http://www.bewitched.com/historyflow.html) より引用した。

えるほど、その編集者は多くの記述を行うようになるが、記事あたりの編集者数が増えると、その編集者は記述の量が少なくなることが分かっている。

De la Gabriel [29] は、Wikipedia の記事には主に 2 つの状況：(1) “stabilized” と呼ばれる、記事が順調に成長している状況、(2) “controversial” と呼ばれる、記事が荒らされていたり議論の最中の状況があると考えた。そこで、状況に応じて質の測定に必要な特徴を変更することによって、質の測定精度を向上させる手法の提案を行っている。

Wikipedia には記事の編集履歴が保存されているため、その情報を利用して記事の状況を推定することができる。Nakamura ら [30] は編集者間の編集行動を解析することによって、編集者のバイアスなど、編集者間の関係や編集者の考え方の違いによって編集者の分類を行う試みが行われている。特に、“尖閣諸島”や“原子力発電所”など、利用者にとって大きく意見が分かれるような記事について、編集者相互の関係を可視化することによって、記事のトピックに関する意見の対立を明らかにすることができる。

3.3 Wikipedia に対する編集者の貢献

編集者が Wikipedia の記事を編集したときや管理を行ったとき、編集者に対して対価が発生しない仕組みとなっている。一方で、Wikipedia に対する編集者の貢献が増加すると、Wikipedia の記事量および質は相対的に向上すると考えられる。Wikipedia へ編集者が記事を編集するうえで考えられる障壁として、テキストを整形する際に必要となるマークアップ言語の習得と、差し戻しなどの編集者にとって否定的な編集が行われたときの編集者のモチベーションの変化が考えられる。そこで、編集者が Wikipedia に対して貢献を行うモチベーションについて解明しようとする研究を紹介する。

編集者が Wikipedia へ貢献を行う際に重要な点として、マークアップ言語に関する知識が乏しい編集者であっても容易に編集を行うことができることがあげられる。Wikipedia では MediaWiki^{*20} において定義されているマークアップ言語およびその拡張言語^{*21} に従って記事を記述する必要がある。もし、このマークアップ言語の習得が非常に難しい場合、編集者は容易に記事の編集を行うことができず、Wikipedia への貢献は困難となってしまうと考えられる。Désilets ら [31] は、数人の子供たちによるグループに Wiki システムを利用させ、マークアップ言語を利用した記事の編集を容易に行うことができることを示した。

編集者が容易に記事を作成することができた場合であっても、記事を編集するモチベーションがなければ編集者が記事を編集することはない。そこで、編集者はどのようなモチベーションによって編集を行ったかについて調査が行われている。Nov [32] は、

Wikipedia の編集者 370 人に対してアンケートを行い、自己の楽しみが主な動機であることを示している。Yang ら [33] が心理学の観点から行った研究では、Leonard ら [34] によって提起されている内的自己概念 (internal self-concept motivation)、つまり自分が思い描く「状況のあるべき姿」の実現を求める傾向が動機として比較的強いが、外的自己概念 (external self-concept motivation)、つまり外部 (世界) が思い描く状況のあるべき姿を実現する傾向と執筆の動機とは、相関関係が見られなかったことを示している。

記事を編集する動機に関しては、初心者である編集者のモチベーションが低下するような場面に遭遇したとき、その後どのような編集行動をとるかについての研究が行われている。Halfaker ら [35] は、編集における差し戻し (Revert) に着目し、差し戻しが利用者の編集に対するモチベーションに対してどのような影響を与えるかを調査している。編集者をあまり編集しない初心者と、頻繁に編集を行う熟練者の 2 つに分け、差し戻しを行ったときに質的な変化、および量的な変化が両者でどのように変化しているかを調査している。ここで、編集者の質を測定する根拠として、4.3.3 項において紹介している研究において用いられることの多い、記述の残存率を利用している。その結果、熟練者が行った編集に対して差し戻しが行われた場合でも編集回数の変化が少ない点に対して、初心者が行った編集に対して差し戻しが行われると、その後その編集者はほとんど編集を行わなくなってしまうことが分かっている。初心者による編集数が減少することは、将来記述されるであろう Wikipedia の記事量が減少することでもありと考えられるため、初心者に対するモチベーションをどのように行うかは課題である。

これらの調査は英語版で行われているが、インターネット上の情報源へ貢献を行う動機は国や文化に依存すると考えられるため、ある言語版で行われた知見が他の言語版においても適用できるとは限らない点に注意すべきである。Ishii ら [36] は、日本と韓国におけるオンラインコミュニケーションの動機が異なることを示している。つまり各言語版において貢献の動機が異なることが予想される。ところが、英語版以外の Wikipedia における記述のモチベーションについての研究を、我々は発見することはできなかった。英語版以外の言語版を対象に編集者の貢献を行う動機を調査した場合、Nov や Yang ら、Halfaker らの知見と同一の知見が得られるとは限らないため、これらの調査が今後行われることが期待される。

Wikipedia は言語版ごとにある程度独立に運営されていること、記述を行っている編集者は言語版に応じて大部分が異なると考えられるため、編集者がなぜ Wikipedia へ貢献を行っているかは各言語版によって異なると考えられる。ところが我々が知る限り、言語版の違いと編集者の貢献度合いとの関係についての調査は存在しない。そのため、これらの調査を行うことによって、編集者の記事への貢献が

^{*20} MediaWiki: <http://www.mediawiki.org/>

^{*21} 目次マークアップ: <http://ja.wikipedia.org/wiki/Help:目次-マークアップ>

言語版によってどの程度違い、記事の質を高めるためにはどのような方針が必要であるかを発見することができると考えられる。

3.4 ガバナンス

Wikipedia は不特定多数の利用者によって記事が作成されているが、無秩序に記事が作成されてしまうと利用者にとって有益ではない情報が大量に作成されてしまう場合がある [37]。ところが、日本語版では、2013 年 3 月現在の管理者は 56 人、活動中の編集者数は 12,072 人、英語版 Wikipedia では管理者が 1,450 人、活動中の編集者数は 135,948 人となっており、編集者数と比較して管理者数は非常に少ない。このような状況下では、管理者は編集者に対して、もしくは編集者相互においてどのような統治（ガバナンス）を行っているかについて、調査が行われている。

Wikipedia では、編集の結果作成された版（バージョン）は、原則としてすべて編集履歴として格納されているが、著作権に違反している記述や個人への誹謗中傷など、編集履歴であっても記述を残すべきでないと考えられる記述については、記事からだけではなく編集履歴からも消去される場合がある。このような懲罰的な処置を公平に行うためには、不適切な記述とは何かを定義する必要があるが、Wikipedia では全体の方針については方針とガイドライン*22において、各々の記事における方針についてはノートページ*23においてそれぞれ望ましい記述、行ってはならない記述について述べられている。そのため、Wikipedia のガバナンスを確立するためには、これらのガイドラインを適切に設定するための過程を構築しなければならない。Forte ら [38] は、Wikipedia の管理に関わる利用者に対してアンケートを行い、ガイドラインの構築過程や秀逸な記事の選択過程などについて明らかにしている。Viégas ら [39] は、協調した編集や方針の遵守にノートページが大きな役割を果たしていることを明らかにしている。また Joo ら [40] は、ノートページから編集者の議論能力を測定する手法を提案している。このように、Wikipedia の管理者へのアンケートやノートページの解析により、Wikipedia のガイドラインを設定する過程が明らかになりつつある。

West ら [41] は、Wikipedia に書かれた不適切な記述が削除されるまでにどの程度の時間が経過しているのか、不適切な記述が何人程度の目に触れているのかに関する調査を行っている。その結果、不適切な記述が行われてから平均 2 分程度で削除が行われており、平均 2 時間以内には編集履歴からも削除が行われていることが分かる。利用者がこのような不適切な記述を閲覧する割合は 0.007% である。不適切である理由の多くは、著作権違反によるものである。

Wikipedia には様々な方針*24があり、それらに従って記事が作成されることが多い。ところが、どのような方針が質の向上に貢献しているかを定量的に測定した研究は、我々が知る限り存在しない。方針は日々更新されているが、方針の変化と記事の質には相関関係があるかどうかは不明確であるため、この相関関係が明らかになることが期待される。

4. 質の測定手法

この章では、Wikipedia における情報の質を測定する方法について紹介する。我々は、主に 2 つの方法によって測定法の分類を行うことができると考えた。1 つの分類法は、測定を行う主体による分類法である。つまり、利用者が質の測定にどの程度介入するかによって分類を行う方法である。もう 1 つは、質の測定を行う指標による分類法である。質の測定を行う主体によって用いることができる指標は異なるため、本論文では測定を行う主体による分類を軸にまず研究の分類を行い、各分類の中で質の測定を行う指標による分類を行う。分類は以下のようになっている。

- 人手による測定 (4.1 節)
 - 人手による測定では、2 章で述べた *trustworthiness* および *expertise* が指標として用いられている。
 - 正確さ (*trustworthiness*) : [8], [42]
 - 専門度 (*expertise*) : [43]
 - 半自動的な測定 (4.2 節)
 - 半自動的な測定では、Wikipedia の閲覧者による投票が指標として用いられている。
 - 投票 : [44], [45], [46], [47]
 - 悪意のある投票の判定 : [48], [49], [50]
 - 自動的な測定 (4.3 節)
 - 自動的な測定では、編集者の編集行動や記事の編集状況が指標として用いられている。1 人の編集者、1 つの記事から得られる指標として編集内容や記事間リンクなどが利用されているが、複数の編集者における相互の関係である編集者間の相互評価なども指標として利用されている。
 - 編集内容 : 単語数 [51], 編集回数・編集者数 [52], [53], 差し戻し回数 [39], 記事の安定性 [54], 編集者間の対立 [55], 文体 [56], 単語の追加・削除 [57], [58]
 - 記事間リンク : リンク数, 被リンク数 [59], [60]
 - 編集者間の相互評価 : [47], [61], [62], [63], [64], [65], [66], [67], [68], [69], [70], [71]
- 以下、それぞれの分類に沿って研究を紹介する。

4.1 人手による測定手法

Wikipedia における記事数は膨大であり、継続的に編集

*22 <http://ja.wikipedia.org/wiki/Wikipedia:方針とガイドライン>

*23 <http://ja.wikipedia.org/wiki/Help:ノートページ>

*24 <http://ja.wikipedia.org/wiki/Wikipedia:方針とガイドラインの一覧>

が行われているため、すべての記事を精査して質を測定することは困難である。そこで、一部の記事に対して人手により質を評価することによって、Wikipedia の記事の質を調査しようとする試みが行われている。本論文では、2章で述べた文献 [14], [15] の分類に従い、trustworthiness と expertise, つまり正確さと専門度の2つの観点によって研究を分類した。

4.1.1 正確さの観点による手法

Giles [8] による 2005 年の調査において、Wikipedia はブリタニカ国際大百科事典^{*25}と比較して信頼度に大きな差異がないことを、Nature の編集者により目視で一部の記事を精査することによって確かめている。この論文によると、Wikipedia には一般の百科事典よりも質の高い記事が数多く存在する一方で、未完成で質の低い記事もまた数多く存在していることが明らかにされている。この調査は反響が大きく、多くの報道機関によって Wikipedia の記事とブリタニカ国際百科事典の記事には正確さに大きな差がなく、十分質が高いと報じられた^{*26}ため、ブリタニカ社では、この論文は不正確であるとの反論 [42] を行っている。これらの議論を受け、Wikipedia を主催する Wikimedia 財団では、アラビア語版、スペイン語版、および英語版の Wikipedia の記事をそれぞれ 6~8 個取り出し、正確さや質に関する詳細なレポートを作成している^{*27}。

4.1.2 専門度の観点による手法

Chesney [43] は、記事の質のうち専門度 (expertise) に関する調査を行っている。研究者を2つのグループに分け、1つのチームにはその研究者の専門に関する記事、もう1つのチームには無作為に選択された記事を評価対象として与えた。つまり、1つのグループは専門家による評価、もう1つのグループは専門家でない利用者による評価を行ったことになる。それぞれの利用者は与えられた記事に対して、編集者と記事そのものの2つを対象とした信憑性の評価を行った。編集者に対する評価は2つのグループにおいて有意な差がなかったが、記事に対する評価は有意差があり、専門家は Wikipedia の記事を信憑性が高いと評価する傾向にあった。この実験の結果から、Wikipedia の記事は比較的信頼できると考えられる一方で、専門家グループに与えられた記事のうち 13% で誤りが見つかったことから、完全に信頼できるわけではないことも明らかにしている。

4.1.3 利点と課題

Wikipedia の記事は多くの利用者によって質の評価がなされているが、4.1 節で述べたものは記事の一部を対象としたものであり、記事全体に対する評価ではない。そのため、

記事の内容を詳細に議論することができるという利点がある一方で、どの記事を評価したかによって評価が大きく変わるという欠点もある。また、2章で述べたように、利用者による質の評価は主観的であり、必ずしも定量的な測定ができるものではない。そこで次に、より客観的に質を評価する手法である半自動的な測定手法について紹介する。

4.2 半自動的な測定手法

半自動的な質の測定方法とは、利用者が記事や記述に対して評価を行い、システムはその利用者の評価から記事に対して質を測定する方法である。主に投票による手法を指すが、投票における問題、および改善手法についても述べる。半自動的な測定手法と自動的な測定手法を組み合わせた手法も存在する [47] が、この手法は 4.3.1 項において述べる。

4.2.1 投票による方法

投票による質の測定手法は情報だけではなく、商品やサービスなど様々な対象に対して質を測定するために、多くのシステムで一般的に利用されている。たとえばショッピングサイトの1つである Amazon.com^{*28}では、利用者は商品に対して5段階の評価を明示的に行っている。システムはこれら利用者が示した評価の平均を閲覧者に示すため、閲覧者は利用者のうちどの程度商品に対して満足しているかを判断することができる。ショッピングサイトだけではなく YouTube^{*29}や Google +1 ボタン^{*30}など多くの Web サービスに実装され、運用されているが、その理由の1つとして簡単に実装できることがあげられる。

Wikipedia においても、同様の機能を実現しようという動きがある。英語版の Wikipedia では Article Feedback Tool^{*31}として、この機能が実装されている。Mizzaro [44] は、論文の査読を自動的に行う方法について述べ、Cusinato ら [45] は、この方法を Wikipedia へ適用する方法について提案している。Sabel ら [46] は、WikiRep と呼ばれるシステムを構築し、利用者からのフィードバックを版と編集者の両方に反映することによって、記事の質を測定している。

誰でも行うことができる投票では、悪意のある投票によって質の測定精度が大きく低下することが考えられる。そこで、悪意のある投票であるかどうかを判定する研究が行われている。Ott ら [48], [49] は、利用者の評価に含まれるコメント文から、どの評価がスパムであるかを特定する手法を提案しており、Mukherjee ら [50] はスパムを行う利用者の評価サイトにおける振舞いを分類することによって、スパムを行う利用者の特定を行う方法を提案している。

^{*25} <http://www.britannica.com/>

^{*26} 「Wikipedia の情報はブリタニカと同じくらい正確」-Nature 誌が調査結果を公表、<http://japan.cnet.com/news/media/20093147/>

^{*27} http://meta.wikimedia.org/wiki/Research:Accuracy_and_quality_of_Wikipedia_entries

^{*28} <http://www.amazon.com/>

^{*29} YouTube: <http://www.youtube.com/>

^{*30} Google +1 ボタン: <http://www.google.com/intl/ja/+1/button/>

^{*31} Article Feedback Tool: http://en.wikipedia.org/wiki/Wikipedia:Article_Feedback_Tool

4.2.2 利点と課題

半自動的な測定手法は、利用者が直感的に正しいと考えられる質の測定を行うことができる点、質の値を測定する根拠が明確である点、システム管理者は利用者の評価入力フォームと評価の表示部分、および集計部分を作成するだけで実装を完了させることができるため、実装が容易であるという点も利点としてあげることができる。

投票による手法は、正確な質を測定することが困難であることも指摘されている。その原因は2つあげることができる。1つは利用者が必ずしも正当な評価を行うわけではないことである。YouTubeと呼ばれる動画投稿サイトにおける Singer による調査^{*32}では、利用者が5段階で動画を評価したときに、ほとんどの評価が5（最高点）であり、1（最低点）が少量であり、2から4の評価がほとんど存在しないことを示している。Wikipedia の Article Feedback Tool においても、90.9%の投票で“useful”（最高点）となっている。これは、78.6%の記事がまだ作成途中であることを示す“B-class”以下の状態であることから、質を測定する対象とその評価には差異があると推定できる。もう1つの原因は4.2.1項でも述べたように、投票に悪意があるかどうかを判定することが困難であることである。たとえば、虚偽の評価に実際に飲食業者に対する評価サイトである「食べログ」においても、虚偽の評価が行われていたことが分かり^{*33}社会問題化している^{*34}。クラウドソーシングに関する研究分野では、判定者の質を測定する手法についても研究されている [72], [73] ため、これらの研究成果を用いて記事の質測定手法を構築することが期待される。

4.3 自動的な測定手法

この節では、利用者の入力を必要とせずにデータだけを利用して情報の質を算出する方法について述べる。

4.3.1 編集内容による方法

質を測定するための指標として、編集者の編集内容から得られる様々な統計量が用いられている。たとえば Blumenstock [51] は単語数、Lih [52] や Wilkinson ら [53] は編集回数や編集者数、Viégas ら [39] は差し戻し回数、Dondio ら [54] は記事の安定性^{*35}、Kittur ら [55] は編集における編集者間の対立などを抽出することによって質を測定している。Kramer ら [47] は、記述の残存率と利用者による判定の2つを組み合わせている。このように、基本的には単語や文字の増減などが特徴として用いられることが多い。

^{*32} Five Stars Dominate Ratings: <http://youtube-global.blogspot.jp/2009/09/five-stars-dominate-ratings.html>

^{*33} 本日の報道内容に関して: <http://corporate.kakaku.com/press/release/20120105.html>

^{*34} 「インターネット消費者取引に係る広告表示に関する景品表示法上の問題点及び留意事項」の一部改定について: http://www.caa.go.jp/representation/pdf/120509premiums_1.pdf

^{*35} 記述の変更が少ない記事は質が高いとする考え方。

統計量だけではなく、自然言語処理などを利用したより複雑な特徴を抽出する試みもある。紙の百科事典は質や文体がある程度均一である点と比較して、Wikipedia の記事では文体が必ずしも均一ではなく、質も一定ではない。このとき、Wikipedia においては文体と質に相関関係があり、硬い文体で書かれた記事は質が高いという考え方がある。Emigh ら [56] は記事をジャンルに分類し、それぞれのジャンルにおけるコーパスを構築することによって、文体の硬さ、柔らかさを測定し、文体の硬さと質に相関関係があることを明らかにしている。

これらの研究においては、質を1つの値として測定することを目標としているが、Stvilja ら [10] は上で述べたような19個の単純な統計値を統合することによって、1つの値を測定することを試みている。ところが実験結果からは、測定された質の値に関する精度そのものよりも、統合の際に必要なパラメータの値が統合後の値の精度に影響を与えていることが判明している。そのため、質を測定する根拠となる特徴量を複数種類抽出する手法を構築する際には、特徴量が質を測定するためにどの程度有益であるかを測定するだけでなく、その特徴量をどのように統合するかについても同時に考慮する必要があると考えられる。

編集者間の関係、版の間の関係を編集内容から抽出して利用することによって、質を測定する方法が考案されている。Brandes ら [57] は、編集者の編集ネットワークに対して最大カット問題を適用することによって、質の高い記事と低い記事へ分類を行う方法について提案している。Zeng ら [58] は、動的ベイジアンネットワーク (Dynamic Bayesian Network) [74] を利用して質を測定する方法を提案している。この手法では、1つのバージョンにおける記事の削除や追加を状態グラフとした状態空間モデルを作成することによって、確率推論によって質の測定を行っている。

編集内容からコミュニティ解析を行う際に、生物工学における成果を活用する方法がある。Dorigo ら [75] は蟻コロニー最適化 (Ant Colony Optimization; ACO) と呼ばれる、アリが群れから食物までの経路を見つける際の挙動からヒントを得た、巡回セールスマン問題に対して近似解を求めるためのアルゴリズムを提案している。Banerjee ら [76] は、Dorigo らが提案した ACO を利用してコミュニティ解析を行うことによって、質の高い記事や低い記事を特定する方法を提案している。

4.3.2 記事間リンクによる方法

相互評価 (Peer review) とは、編集者相互や記事相互で評価し、質の測定を行う方法である。Wikipedia では、編集者が直接他の編集者に対して評価を行わない。そのため、編集者の編集活動から評価を推測する必要がある。このような間接的な評価は Wikipedia 以外にも様々な分野で行われてきた古典的な問題であり、様々な解決方法が提案されてきた。

たとえば学術界においても、論文の評価において間接的な評価を行うことがあり、有名なものとしてインパクトファクタ [77] がある。この指標は、学術雑誌に対して質が高いかどうかを計測するための方法であり、多く引用される論文は良いという仮定に基づいている。つまり、引用という論文編集者の活動から、評価を推測していることになる。Web のリンク解析に用いられる HITS [78] や PageRank [79], SALSA [80] では、論文の引用に相当する関係として Web ページ間のハイパーリンクを利用して、リンク数および被リンク数が多い Web ページは質が高いという仮定に基づいた Web ページの評価を行っている。Wikipedia においても同様に、Bellomi ら [59] や Wu ら [60] によって Wikipedia の内部リンクを利用して HITS や PageRank によって質の算出を行っている。これらの手法は一定の成果をおさめているが、Wikipedia においてはこれらの手法は必ずしも有効であるとはいえない。なぜならば、リンク数、被リンク数が多い Web ページの質が高いという仮定が必ずしも正しいとはいえず、質が低く十分な記述がない記事などへもリンクが多数作成されることがあるためである。Bellomi らの評価実験では、一般名詞やよく利用される単語に関する記事は質が高いと判定されているが、秀逸な記事などは質が高いと判定されていない。

4.3.3 残存率による方法

編集者間の相互評価による質の測定手法も提案されている。Adler ら [62], [63], [64], Hu ら [70] は、編集履歴を用いて質の測定を行う方法を提案している。Wikipedia において記述が行われたとき、その記述の質が高いときには他の編集者から削除されないが、質が低いときには削除されることが多い。この仮定は、3.3 節で述べた West ら [41] の実験からも成立する場合が多いことが期待できる。つまり、多くの編集を経て残存する記述は質が高いと推定することができる。そこで、編集履歴を用いて記述の編集における残存率を計算することによって、質の計算を行っている。この手法では、Wikipedia の記事全体の質の平均が編集経過とともに上昇することを仮定している。

この手法には 2 つの問題点がある。1 つは計算コストが大きいこと、もう 1 つは悪質な利用者からの攻撃に弱いことである。1 つ目の問題を解決するために、鈴木ら [69] はキーパーソン抽出による計算コスト削減を提案している。Wikipedia において、多くの編集者は記述する回数が少なく、少数の著者が大部分の記述を行っている。そこで、すべての編集者による記述ではなく、これら少数の編集者だけに対して質を測定し、その編集者による記述の質を編集者の質で代用することができれば、計算時間を削減できると考えた。実際に、30% から 40% の計算時間で Adler らの手法による質の測定精度とほぼ変わらない精度を実現することができている。

もう 1 つの問題である悪質な編集者からの攻撃には様々

なものと考えられるが、1 つは 1 人の利用者が多重アカウント (ソックパペット^{*36}) を用いて、不当に記述の残存率を変更する攻撃 (Sybil Attack)、もう 1 つは質の高い記述を削除し、質の低い記述を残存させるという攻撃によって、Adler らの手法における編集経過とともに質が向上するという仮定が必ずしも成立しない場合に、質の測定精度が低下するという問題 (Zig-zag attack) である。1 つ目の攻撃については Chatterjee ら [71] によって解決手法が提案されている。この手法では、編集者がテキストを記事に挿入するとき、通常の編集者であれば 1 度の編集で行うべきものを複数回の編集によって行い、かつその複数回の編集が 1 人の編集者によって行われる場合には、その編集を不当に残存率を向上させる行為であるとして検出する。さらに、その複数回の編集が複数の編集者群によって行われたときには、その編集者群は多重アカウントの可能性が高いと検出する。

もう 1 つの攻撃である zig-zag attack については、Suzuki ら [66], [68] によって解決手法が提案されている。この手法では、特に削除行為に着目し、編集を行う編集者と記述を削除された編集者を 2 つのグループとし、それぞれの編集者をノードとした二部グラフを構築する。このグラフに対して HITS アルゴリズムのように編集者と記述を相互に評価し、良質な編集者、記述を測定する。このとき、質の高い編集者は質の高い記事を残し、質の低い記事を削除する、質の低い編集者は質の高い記事を削除し、質の低い記事を残す傾向にあることを仮定している。この仮定から質の高い編集者を特定し、編集者の質の高さによって他の編集者の質への影響の強さを変更した。このとき、悪質な編集者は他の編集者への質への影響の強さが小さいため、攻撃を行うことによって意図的に質の値を変更することが困難となる。

残存率による質の測定手法では、記述が残存している割合が大きければ大きいほど質が高いと判定される。このとき、多くの手法では版ごとの差異を測定するために編集距離 (Edit Distance) を利用することが多い。つまり、編集における追加と削除は同一の重みとして扱われている。ところが、実際には追加と削除には大きな差があり、異なる重みを与えることによって質の測定精度が向上すると考えられる。Suzuki ら [65], [67] は、編集者が記述を残存させることを肯定的な編集 (positive ratings)、削除することを否定的な編集 (negative ratings) としてそれぞれ個別に扱うことによって、質の測定精度を改善している。

4.3.4 利点と課題

4.3 節で述べている自動的な質の測定手法は、Wikipedia で公開されているデータだけを利用して質を測定する手法であるため、客観的な質の測定を行うことが可能である点

^{*36} 多重アカウント : <http://ja.wikipedia.org/wiki/Wikipedia:多重アカウント>

は利点であるといえる。ところが、4.2節で述べた半自動的な質の測定手法と異なり、質の高さを利用者が明示しない方法であるため、人間による主観の利用が困難である点は欠点である。2章で述べたように、情報の質は人間の主観による度合いであるにもかかわらず、この章で述べる方法では利用者の入力を用いないため、利用者の直感を直接反映させた質の判定を行うことは難しい。

この問題を解決する方法の1つとして、文体を用いる方法が考えられる。Wikipediaのテキストは電子掲示板などの記述と比較して硬い文体であることが多いため、比較的意味解析を行うことが容易であると考えられる。ただし、意味解析結果を用いることによってどのように質を測定することができるかは明らかではない。記事に対して意味解析を行うことによって、より高度な質の測定手法を確立することが必要である。

さらに、質の算出単位についても考慮する必要がある。質を算出する単位としては、記述の一部、記述の全体、および編集者などが主に利用されているが、それぞれ問題が考えられる。記述の一部や全体を質の算出単位としたときには、記述がどのように変化したかが主な方法となり、新しい記述に対して質を測定できないという問題がある。また、編集者を質の算出単位としたときには、1人で複数のアカウントを利用している場合や、複数人でWikipediaに対して攻撃を行ったときに対応することが困難である。そのため、質を測定する単位も重要な問題である。

5. おわりに

本論文では、Wikipediaにおける記事の状況、および質の測定手法に関する研究を紹介した。現在、インターネット上に存在する数多くの情報の中で、Wikipediaは比較的質の高い情報であると認識されていることが多いが、Wikipediaの記事を編集することへの参入障壁が比較的低いため、質が低い記事が存在することも確かである。質とは2章でも述べたとおり複雑な概念であり、質を測定するためには、単純に真実や事実であるかどうかの判定に限らず、多くの人がある情報をどのように感じているかなど、様々な要因を考慮して決定しなければならない。このように、質の測定は非常に困難な課題であるにもかかわらず、現在多くの研究者によって取り組まれている課題である。

我々は、情報の質を測定する手法、および質の測定を補助する方法は今後ますます重要となると考えている。Wikipediaには、人手ですべての記事に対して質を測定することが困難であるほどの大量の情報存在する。その一方で、情報の質を測定するためには、その情報に関する十分な知識が必要である。知識が十分でない利用者が質の低い情報に接したとき、その利用者が不利益を被ったり、流言が発生し社会的に大きな被害が発生する可能性もある。このような不利益や被害を防止するためにも、大量の情報

に対して質を測定する方法に対する取り組みは重要であると考えられる。

Wikipediaの質を測定する手法と他言語版記事の内容を活用することによって、記事の量を増加させることも可能となる。Nadamotoら[81],[82]は、ある言語版において言及されているが他の言語版では言及されていない内容を補完する手法についての提案を行っている。この手法を利用することによって、内容の少ない言語版の記事に対して、内容が多く質の高い言語版の記事内容を補完することができる。

また、Wikipediaにおける情報の質を測定し改善することによって、今後様々な情報源の質を改善することが期待できる。文献[5]で述べているように、Wikipediaを情報源として活用するアプリケーションとして、コーパスの作成や知識ベースの構築、語義曖昧性解消、情報検索など様々なものが存在する。これらのアプリケーションではWikipediaに含まれる情報の質が高いことを前提としたものが多いため、質が低い記事を扱った場合に精度が低下することが予想される。Wikipediaから生成された知識ベースであるYAGO2[83]では、この問題を解決するためにWikipedia以外のデータソースを併用しているが、それぞれのデータソースの質が十分でなければ、質の高い知識ベースを構築することは困難である。そのため、これらのアプリケーションに対して記事の質測定手法を組み込むことによって、精度が向上し利便性の高いアプリケーションとなることが期待される。

謝辞 本研究の一部は、JSPS 科研費 23700113 の助成を受けたものです。ここに記して謝意を表します。

参考文献

- [1] Voss, J.: Measuring Wikipedia, *Proc. Int'l Conference of the Int'l Society for Scientometrics and Informetrics: 10th*, pp.24-28 (2005) (online), available from <http://eprints.rclis.org/archive/00003610/>.
- [2] Remy, M.: Wikipedia: The Free Encyclopedia, *Online Information Review*, Vol.26, No.6, pp.434-435 (2002).
- [3] 日下九八: ウィキペディア: その信頼性と社会的役割, *情報管理*, Vol.55, No.1, pp.2-12 (2012).
- [4] 渡辺智暁: われわれはウィキペディアとどうつきあうべきか: メディア・リテラシーの視点から (特集) ソーシャルサービス活用指南, *情報の科学と技術*, Vol.61, No.2, pp.64-69 (2011).
- [5] 中山浩太郎, 伊藤雅弘, Maiké, E., 白川真澄, 道下智之, 原隆浩, 西尾章治郎: Wikipedia マイニング: Wikipedia 研究のサーベイ, *情報処理学会論文誌データベース (TOD)*, Vol.2, No.4, pp.49-60 (2009).
- [6] 松林正己, 白石 啓, 権田真幸, 高久雅生: 特集: 「情報の信頼性」の編集にあたって, *情報の科学と技術*, Vol.61, No.1, p.1 (2011).
- [7] 大島裕明, 山本祐輔, 山家雄介, 高橋良平, ヤトフトアダム, 中村聡史, 田中克己: Web 情報の信憑性 (特集) 情報の信頼性, *情報の科学と技術*, Vol.61, No.1, pp.2-7 (2011).
- [8] Giles, J.: Special report: Internet Encyclopedias Go Head to Head, *Nature*, Vol.438, No.15, pp.900-901

- (2005).
- [9] Denning, P., Horning, J., Parnas, D. and Weinstein, L.: Wikipedia Risks, *Comm. ACM*, Vol.48, No.12, p.152 (online), DOI: 10.1145/1101779.1101804 (2005).
- [10] Stvilia, B., Twidale, M.B., Smith, L.C. and Gasser, L.: Information Quality Work Organization in Wikipedia, *J. Am. Soc. Inf. Sci. Technol.*, Vol.59, No.6, pp.983–1001 (online), DOI: 10.1002/asi.v59:6 (2008).
- [11] Wand, Y. and Wang, R.Y.: Anchoring Data Quality Dimensions in Ontological Foundations, *Comm. ACM*, Vol.39, No.11, pp.86–95 (online), DOI: 10.1145/240455.240479 (1996).
- [12] Wang, R.Y., Storey, V.C. and Firth, C.P.: A Framework for Analysis of Data Quality Research, *IEEE Trans. Knowl. and Data Eng.*, Vol.7, No.4, pp.623–640 (online), DOI: 10.1109/69.404034 (1995).
- [13] 新村 出, 新村出記念財団: 広辞苑 第 6 版 (2009).
- [14] Hovland, C.I., Janis, I.L. and Kelley, H.H.: *Communication and Persuasion; Psychological Studies of Opinion Change*, Yale University Press (1953).
- [15] Fogg, B.J. and Tseng, H.: The Elements of Computer Credibility, *Proc. SIGCHI Conference on Human Factors in Computing Systems: The CHI is the Limit, CHI '99*, pp.80–87, ACM (online), DOI: 10.1145/302979.303001 (1999).
- [16] Metzger, M.J., Flanagin, A.J., Eyal, K., Lemus, D. and McCann, R.: Credibility for the 21st Century: Integrating Perspectives on Source, Message, and Media Credibility in the Contemporary Media Environment, *Communication Yearbook*, Vol.27, pp.293–335 (2003).
- [17] Wong, M.A.: *Reference and Information Services: An Introduction, 3rd edition*, Chapter Encyclopedias, pp.433–459, Englewood (2011).
- [18] Sheppard, S.A. and Terveen, L.: Quality is a Verb: The Operationalization of Data Quality in a Citizen Science Community, *Proc. 7th Int'l Symposium on Wikis and Open Collaboration, WikiSym '11*, pp.29–38, ACM Press (online), DOI: 10.1145/2038558.2038565 (2011).
- [19] White, D.: Results and Analysis of Web 2.0 Services Survey Undertaken by the SPIRE Project (2007).
- [20] Flanagin, A. and Metzger, M.: FROM ENCYCLOPEDIA BRITANNICA TO WIKIPEDIA, *Information, Communication & Society*, Vol.14, No.3, pp.355–374 (online), DOI: 10.1080/1369118X.2010.542823 (2011).
- [21] Holloway, T., Bozicevic, M. and Börner, K.: Analyzing and Visualizing the Semantic Coverage of Wikipedia and its Authors, *Complexity*, Vol.12, No.3, pp.30–40 (2007).
- [22] Ortega, F. and Gonzalez-Barahona, J.M.: Quantitative Analysis of the Wikipedia Community of Users, *Proc. 2007 Int'l Symposium on Wikis, WikiSym '07*, pp.75–86, ACM Press (online), DOI: 10.1145/1296951.1296960 (2007).
- [23] Kittur, A., Suh, B. and Chi, E.H.: Can You ever Trust a Wiki?: Impacting Perceived Trustworthiness in Wikipedia, *Proc. Conference on Computer Supported Cooperative Work (CSCW '08)*, pp.477–480, ACM Press (online), DOI: <http://doi.acm.org/10.1145/1460563.1460639> (2008).
- [24] Viégas, F.B., Wattenberg, M. and Dave, K.: Studying Cooperation and Conflict between Authors with *History Flow* Visualizations, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI '04*, pp.575–582, ACM Press (2004).
- [25] Sabel, M.: Structuring Wiki Revision History, *Proc. 2007 Int'l Symposium on Wikis WikiSym '07*, pp.125–130, ACM Press (online), DOI: 10.1145/1296951.1296965 (2007).
- [26] Deza, M. and Deza, E.: *Encyclopedia of Distances, 1st edition*, Springer (2009).
- [27] Roth, C., Taraborelli, D. and Gilbert, N.: Measuring Wiki Viability: An Empirical Assessment of the Social Dynamics of a Large Sample of Wikis, *Proc. 4th Int'l Symposium on Wikis, WikiSym '08*, pp.27:1–27:5, ACM Press (online), DOI: 10.1145/1822258.1822294 (2008).
- [28] Roth, C. and Cointet, J.-P.: Social and Semantic Coevolution in Knowledge Networks, *Social Networks*, Vol.32, No.1, pp.16–29 (online), DOI: 10.1016/j.socnet.2009.04.005 (2010).
- [29] De la Calzada, G. and Dekhtyar, A.: On Measuring the Quality of Wikipedia Articles, *Proc. 4th Workshop on Information Credibility, WICOW '10*, pp.11–18, ACM Press (online), DOI: 10.1145/1772938.1772943 (2010).
- [30] Nakamura, A., Suzuki, Y. and Ishikawa, Y.: Clustering Editors of Wikipedia by Editor's Biases, *2013 IEEE/WIC/ACM Int'l Conference on Web Intelligence*, IEEE CS Press (2013).
- [31] Désilets, A., Paquet, S. and Vinson, N.G.: Are Wikis Usable?, *Proc. 2005 Int'l Symposium on Wikis, WikiSym '05*, pp.3–15, ACM Press (online), DOI: 10.1145/1104973.1104974 (2005).
- [32] Nov, O.: What Motivates Wikipedians?, *Comm. ACM*, Vol.50, No.11, pp.60–64, ACM Press, DOI: 10.1145/1297797.1297798 (2007).
- [33] Yang, H.-L. and Lai, C.-Y.: Motivations of Wikipedia Content Contributors, *Comput. Hum. Behav.*, Vol.26, No.6, pp.1377–1383 (online), DOI: 10.1016/j.chb.2010.04.011 (2010).
- [34] Leonard, N.H., Beauvais, L.L. and Scholl, R.W.: Work Motivation: The Incorporation of Self-Concept-Based Processes, *Human Relations*, Vol.52, No.8, pp.969–998 (online), DOI: 10.1023/A:1016927507008 (1999).
- [35] Halfaker, A., Kittur, A. and Riedl, J.: Don't Bite the Newbies: How Reverts Affect the Quantity and Quality of Wikipedia Work, *Proc. 7th Int'l Symposium on Wikis and Open Collaboration, WikiSym '11*, pp.163–172, ACM Press (online), DOI: 10.1145/2038558.2038585 (2011).
- [36] Ishii, K. and Ogasahara, M.: Links between Real and Virtual Networks: A Comparative Study of Online Communities in Japan and Korea, *Cyberpsychol Behav*, Vol.10, No.2, pp.252–257 (online), DOI: 10.1089/cpb.2006.9961 (2007).
- [37] Surowiecki, J.: *The Wisdom of Crowds*, Anchor (2005).
- [38] Forte, A. and Bruckman, A.: Scaling Consensus: Increasing Decentralization in Wikipedia Governance, *Proc. 41st Annual Hawaii Int'l Conference on System Sciences, HICSS '08*, p.157, IEEE Computer Society (online), DOI: 10.1109/HICSS.2008.383 (2008).
- [39] Viégas, F.B., Wattenberg, M., Kriss, J. and van Ham, F.: Talk Before You Type: Coordination in Wikipedia, *Proc. 40th Annual Hawaii Int'l Conference on System Sciences, HICSS '07*, p.78, IEEE Computer Society (online), DOI: 10.1109/HICSS.2007.511 (2007).
- [40] Joo, S. and Takeda, H.: Analysis of Discussion Page in Wikipedia Based on User's Discussion Capability, *Proc. 2012 IEEE/WIC/ACM Int'l Joint Conferences on Web Intelligence and Intelligent Agent Technology - Volume 01, WI-IAT '12*, pp.243–247, IEEE Computer Society (online) (2012).

- [41] West, A.G. and Lee, I.: What Wikipedia Deletes: Characterizing Dangerous Collaborative Content, *Proc. 7th Int'l Symposium on Wikis and Open Collaboration, WikiSym '11*, pp.25–28, ACM Press (online), DOI: 10.1145/2038558.2038563 (2011).
- [42] Encyclopaedia Britannica: Fatally flawed. Refuting the Recent Study on Encyclopedic Accuracy by the Journal Nature (2006).
- [43] Chesney, T.: An Empirical Examination of Wikipedia's Credibility, *First Monday*, Vol.11, No.11 (2006) (online).
- [44] Mizzaro, S.: Quality Control in Scholarly Publishing., *J. Am. Soc. Inf. Sci. Technol.*, Vol.54, pp.989–1005 (online), DOI: 10.1.1.9.1961 (2003).
- [45] Cusinato, A., D. Mea, V., D. Salvatore, F. and Mizzaro, S.: QuWi: Quality Control in Wikipedia, *Proc. 3rd Workshop on Information Credibility on the Web, WICOW '09*, pp.27–34, DOI: <http://doi.acm.org/10.1145/1526993.1527001> (2009).
- [46] Sabel, M., Garg, A. and Battiti, R.: WikiRep: Digital Reputations in Virtual Communities, *Proc. XLIII Congresso Annuale AICA*, pp.209–217 (2005).
- [47] Kramer, M., Gregorowicz, A. and Iyer, B.: Wiki Trust Metrics Based on Phrasal Analysis, *Proc. 4th Int'l Symposium on Wikis, WikiSym '08*, pp.24:1–24:10, ACM Press (online), DOI: 10.1145/1822258.1822291 (2008).
- [48] Ott, M., Choi, Y., Cardie, C. and Hancock, J.T.: Finding Deceptive Opinion Spam by Any Stretch of the Imagination, *Proc. 49th Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1, HLT '11, ACL*, pp.309–319 (online) (2011).
- [49] Ott, M., Cardie, C. and Hancock, J.: Estimating the Prevalence of Deception in Online Review Communities, *Proc. 21st Int'l Conference on World Wide Web, WWW '12*, pp.201–210, ACM Press (online), DOI: 10.1145/2187836.2187864 (2012).
- [50] Mukherjee, A., Liu, B. and Glance, N.: Spotting Fake Reviewer Groups in Consumer Reviews, *Proc. 21st Int'l Conference on World Wide Web, WWW '12*, pp.191–200, ACM Press (online), DOI: 10.1145/2187836.2187863 (2012).
- [51] Blumenstock, J.E.: Size Matters: Word Count as a Measure of Quality on Wikipedia, *Proc. 17th Int'l Conference on World Wide Web, WWW '08*, pp.1095–1096, ACM Press (online), DOI: 10.1145/1367497.1367673 (2008).
- [52] Lih, A.: Wikipedia as Participatory Journalism: Reliable sources? Metrics for Evaluating Collaborative Media as a News Resource, *Proc. 5th Int'l Symposium on Online Journalism*, pp.16–17 (2004).
- [53] Wilkinson, D.M. and Huberman, B.A.: Cooperation and quality in wikipedia, *Proc. 2007 Int'l Symposium on Wikis, WikiSym '07*, pp.157–164, ACM Press (online), DOI: 10.1145/1296951.1296968 (2007).
- [54] Dondio, P., Barrett, S., Weber, S. and Seigneur, J.M.: Extracting Trust from Domain Analysis: A Case Study on the Wikipedia Project, *Proc. 3rd Int'l Conference on Autonomic and Trusted Computing, ATC'06*, pp.362–373, Springer-Verlag (2006).
- [55] Kittur, A., Suh, B., Pendleton, B.A. and Chi, E.H.: He Says, She Says: Conflict and Coordination in Wikipedia, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI '07*, pp.453–462, ACM Press (online), DOI: 10.1145/1240624.1240698 (2007).
- [56] Emigh, W. and Herring, S.: Collaborative Authoring on the Web: A Genre Analysis of Online Encyclopedias, *Proc. 38th Annual Hawaii Int'l Conference on System Sciences, 2005, HICSS '05*, p.99a (online), DOI: 10.1109/HICSS.2005.149 (2005).
- [57] Brandes, U., Kenis, P., Lerner, J. and van Raaij, D.: Network analysis of collaboration structure in Wikipedia, *Proc. 20th Int'l Conference on World Wide Web, WWW '09*, pp.731–740 (2009).
- [58] Zeng, H., Alhossaini, M.A., Ding, L., Fikes, R. and McGuinness, D.L.: Computing Trust from Revision History, *Proc. 2006 Int'l Conference on Privacy, Security and Trust: Bridge the Gap Between PST Technologies and Business Services, PST '06*, p.8:1, ACM Press (online), DOI: 10.1145/1501434.1501445 (2006).
- [59] Bellomi, F. and Bonato, R.: Network Analysis for Wikipedia, *Proc. Wikimania*, (online), DOI: 10.1.1.178:2233 (2005).
- [60] Wu, I.-C. and Wu, C.-Y.: Using Internal Link and Social Network Analysis to Support Searches in Wikipedia: A Model and its Evaluation, *J. Inf. Sci.*, Vol.37, No.2, pp.189–207 (2011).
- [61] De Alfaro, L., Kulshreshtha, A., Pye, I. and Adler, B.T.: Reputation Systems for Open Collaboration, *Comm. ACM*, Vol.54, No.8, pp.81–87, DOI: 10.1145/1978542.1978560 (2011).
- [62] Adler, B.T., de Alfaro, L., Pye, I. and Raman, V.: Measuring Author Contributions to the Wikipedia, *Proc. 4th Int'l Symposium on Wikis, WikiSym '08*, pp.15:1–15:10, ACM Press (online), DOI: 10.1145/1822258.1822279 (2008).
- [63] Adler, B.T., Chatterjee, K., de Alfaro, L., Faella, M., Pye, I. and Raman, V.: Assigning Trust to Wikipedia Content, *Proc. 4th Int'l Symposium on Wikis, WikiSym '08*, pp.26:1–26:12, ACM Press (online), DOI: 10.1145/1822258.1822293 (2008).
- [64] Adler, B. and de Alfaro, L.: A Content-driven Reputation System for the Wikipedia, *Proc. 16th Int'l Conference on World Wide Web, WWW '07*, pp.261–270 (online), DOI: <http://doi.acm.org/10.1145/1242572>. 1242608 (2007).
- [65] Suzuki, Y.: Effects of Implicit Positive Ratings for Quality Assessment of Wikipedia Articles, *Journal of Information Processing*, Vol.21, No.2, pp.342–348 (2013).
- [66] Suzuki, Y. and Yoshikawa, M.: Mutual Evaluation of Editors and Texts for Assessing Quality of Wikipedia Articles, *Proc. Int'l Symposium on Wikis and Open Collaboration, WikiSym '12*, (online), DOI: 10.1145/2462932.2462956 (2012).
- [67] Suzuki, Y.: Assessing Quality Values of Wikipedia Articles Using Implicit Positive and Negative Ratings, *Proc. 13th Int'l Conference on Web-Age Information Management (WAIM 2012)*, pp.127–138 (2012).
- [68] Suzuki, Y. and Yoshikawa, M.: QualityRank: Assessing Quality of Wikipedia Articles by Mutually Evaluating Editors and Text, *Proc. 23rd ACM Conference on Hypertext and Social Media (HT 2012)*, pp.307–308, ACM Press (2012).
- [69] 鈴木 優, 吉川正俊: Wikipedia におけるキーパーソン抽出による信頼度算出精度および速度の改善, 情報処理学会論文誌: データベース, Vol.3, No.3, pp.20–32 (2010).
- [70] Hu, M., Lim, E., Sun, A., Lauw, H.W. and Vuong, B.: Measuring Article Quality in Wikipedia: Models and Evaluation, *Proc. ACM Int'l Conference on Information and Knowledge Management (CIKM 2007)*, pp.243–252, ACM Press (2007).
- [71] Chatterjee, K., de Alfaro, L. and Pye, I.: Robust

- Content-driven Reputation, *AISeC*, pp.33–42 (2008).
- [72] Kittur, A., Chi, E.H. and Suh, B.: Crowdsourcing User Studies with Mechanical Turk, *Proc. SIGCHI Conference on Human Factors in Computing Systems, CHI '08*, pp.453–456, ACM Press (online), DOI: 10.1145/1357054.1357127 (2008).
- [73] Bernstein, M.S., Little, G., Miller, R.C., Hartmann, B., Ackerman, M.S., Karger, D.R., Crowell, D. and Panovich, K.: Soylent: A Word Processor with a Crowd Inside, *Proc. 23rd Annual ACM Symposium on User Interface Software and Technology, UIST '10*, pp.313–322, ACM Press (online), DOI: 10.1145/1866029.1866078 (2010).
- [74] Cowell, R.G., Lauritzen, S.L., David, A.P. and Spiegelhalter, D.J.: *Probabilistic Networks and Expert Systems, 1st edition*, Springer-Verlag (1999).
- [75] Dorigo, M. and Blum, C.: Ant Colony Optimization Theory: A Survey, *Theor. Comput. Sci.*, Vol.344, No.2–3, pp.243–278 (2005).
- [76] Banerjee, S., El-Bendary, N. and Al-Qaheri, H.: Exploring Wiki: Measuring the Quality of Social Media Using Ant Colony Metaphor, *Proc. Int'l Conference on Management of Emergent Digital EcoSystems, MEDES '11*, pp.305–312, ACM Press (online), DOI: 10.1145/2077489.2077545 (2011).
- [77] Garfield, E.: Citation Indexes for Science; a New Dimension in Documentation through Association of Ideas, *Science*, Vol.122, No.3159, pp.108–111 (1955).
- [78] Kleinberg, J.M.: Authoritative Sources in a Hyperlinked Environment, *J. ACM*, Vol.46, pp.604–632, ACM Press, DOI: <http://doi.acm.org/10.1145/324133.324140> (1999).
- [79] Brin, S. and Page, L.: The Anatomy of a Large-Scale Hypertextual Web Search Engine, *Proc. 6th Int'l Conference on World Wide Web, WWW'97*, CDROM (1997).
- [80] Lempel, R. and Moran, S.: SALSA: The Stochastic Approach for Link-structure Analysis, *ACM Trans. Inf. Syst.*, Vol.19, No.2, pp.131–160, ACM Press, DOI: 10.1145/382979.383041 (2001).
- [81] Fujiwara, Y., Suzuki, Y., Konishi, Y. and Nadamoto, A.: Extracting Difference Information from Multilingual Wikipedia, *Proc. 14th Asia-Pacific Int'l Conference on Web Technologies and Applications, APWeb'12*, pp.496–503, Springer-Verlag (2012).
- [82] Suzuki, Y., Fujiwara, Y., Konishi, Y. and Nadamoto, A.: Good Quality Complementary Information for Multilingual Wikipedia, *Proc. 13th Int'l Conference on Web Information Systems Engineering, WISE'12*, pp.185–198, Springer-Verlag (2012).
- [83] Hoffart, J., Suchanek, F.M., Berberich, K. and Weikum, G.: YAGO2: A Spatially and Temporally Enhanced Knowledge Base from Wikipedia, *Artif. Intell.*, Vol.194, pp.28–61, DOI: 10.1016/j.artint.2012.06.001 (2013).



鈴木 優 (正会員)

1977年生。1999年神戸大学工学部情報知能工学科卒業，2001年奈良先端科学技術大学院大学博士前期課程修了。2004年同博士課程修了。博士(工学)。2004年立命館大学講師。2009年京都大学特定研究員。2010年名古屋大学研究員，2011年同大学特任助教。ソーシャルメディアの解析，情報検索の研究に従事。日本データベース学会，電子情報通信学会，人工知能学会，IEEE-CS，ACM各会員。

(担当編集委員 牛尼 剛聡)