How Did They Build the Free Encyclopedia? A Literature Review of Collaboration and Coordination among Wikipedia Editors

YUQING REN, University of Minnesota, USA

HAIFENG ZHANG, Carnegie Mellon University, USA

ROBERT E. KRAUT, Carnegie Mellon University, USA

Wikipedia has been the poster child for large-scale online open collaboration while few other online open collaboration initiatives have achieved similar success. How did Wikipedians do it? Besides the technical infrastructure, what social dynamics and processes are critical to its success? This essay reviews 217 articles that examined aspects of the behaviors of Wikipedia editors and the processes through which they coordinate and collaborate. Using the Input-Mediator-Output-Input model (IMOI) as the organizing framework, we summarized the key insights in an integrative review. The input factors include editors, their motivations, and the tools they use to support their work. The mediating factors include coordination, governance, leadership, conflict, newcomer socialization, and roles. The outcome focuses on measuring and predicting contribution quantity and quality. We hope our work serves as a road map for researchers who are interested in Wikipedia to learn about prior research and identify future research directions.

CCS Concepts: • **General and reference** → **Surveys and overviews**; • **Human-centered computing** → **Computer supported cooperative work**.

Additional Key Words and Phrases: Wikipedia, survey, literature review, editors, collaboration

1 INTRODUCTION

In the past two decades, the Internet has made it possible for millions of individuals across the globe to collaborate and create digital artifacts and knowledge sources. One of the most prominent cases is Wikipedia, the free online encyclopedia. Although there are many similar initiatives such as Zooniverse, Encyclopedia of Life, Project Gutenberg, MusicBrainz, and open source development projects, where netizens collaborate online to create artifacts of lasting value [35], few have achieved similar levels of success as Wikipedia. What are the key factors to Wikipedia's success? Who are the Wikipedians and how do they orchestrate individual efforts to build such a large-scale open collaboration project? What can we learn from Wikipedia, especially the working and collaboration among Wikipedians that will make a new large-scale open collaboration successful? These are the questions we hope to answer in this review article.

Several reasons make understanding the success of Wikipedia important. First, Wikipedia is one of the most used information sources in the world, and understanding what makes it successful is important in its own right. According

Authors' addresses: Yuqing Ren, University of Minnesota, USA, chingren@umn.edu; Haifeng Zhang, Carnegie Mellon University, USA, haifeng1@andrew.cmu.edu; Robert E. Kraut, Carnegie Mellon University, USA, robert.kraut@cmu.edu.

to the web analytics company Similarweb, as of May 2023, Wikipedia was the seventh most visited website in the world [164]. As of January 2023, Wikipedia existed in 292 languages with the English version of Wikipedia featuring 6.6 million articles and over 20 gigabytes of information [187]. Second, an analysis of 35 years of human-computer interaction research showed a shift from studying primarily humans interacting with computers to studying social interactions among human users [80]. Understanding Wikipedia's success can inform research related to collaborative and social computing systems, a relatively newer domain in the field of human-computer interaction [214]. For example, insights related to editors' motivations to contribute, causes of the under-representation of women, and the factors that influence the growth and decline of Wikipedia may generalize to other collaborative and social computing systems. Third, an important goal of human-computer interaction research is to identify best practices and make design recommendations. Wikipedia is a well-designed socio-technical system and the Wikimedia Foundation maintains longitudinal records of the technical and social designs as well as editors' activities and work products. Understanding Wikipedia's success can inform the design of future collaborative and social computing systems.

Since its early days, Wikipedia's exponential growth has stimulated substantial interest from the academic community. We searched for academic articles using the keyword "Wikipedia" and found 5,756 results from Web of Science, 14,305 results from ACM Digital Library, and over 2.27 million results from Google Scholar. Despite the enormous volume of research around Wikipedia, what makes Wikipedia successful remains somewhat of a mystery, primarily due to the difficulty of aggregating and synthesizing results from hundreds of studies.

There have been prior efforts to develop comprehensive reviews of Wikipedia research. For example, Okoli et al. [141] conducted an exhaustive review of research related to Wikipedia and identified six themes around content, corpus, infrastructure, participation, readership, and a miscellaneous general category. Jullien [91] compared studies across three languages – English, French, and Spanish – and summarized research around editor motivations, the processes and patterns of interactions among editors, and the organization of the Wikipedia community. There have also been reviews of specialized topics around content [124], readership [142], vandalism [174], and the use of Wikipedia as a corpus for information retrieval purposes [121, 122]. Almost ten years have passed since the two comprehensive reviews by Jullien [91] and Okoli et al. [141]. Furthermore, a significant amount of work around Wikipedia collaboration has been published in conference proceedings that were not well covered in previous reviews.

In this paper, we hope to expand prior reviews in three ways: adding coverage of recent research especially articles published between 2012 and 2020 (e.g., topics related to bots, gender gap, and newcomer socialization), inclusion of conference proceedings, especially ACM conferences, and an integrative framework to summarize insights around the social and community aspects of Wikipedia. As an evolving socio-technical system, Wikipedia's success depends on the interplay of its technical infrastructure and its community of volunteer editors [91]. [1] As the community grows, the challenge of coordinating the large number of volunteers increases, and researchers have called for processes, policies, and a complex organizational structure to facilitate the large-scale collaboration this challenge demands [6]. Our primary focus in this review is the human behaviors and social processes through which Wikipedia collaboration occurs. By reviewing a representative body of research, we hope to achieve two goals: to draw insights about how Wikipedia collaboration works, which can be applied to other collaboration initiatives, and to identify opportunities for future research.

---

[1]Wikipedia [206] also lists other factors key to Wikipedia's success such as being a public project, online and open access, attracting both amateurs and professionals, editors' autonomy in choosing their tasks, continuous and fast editing, consensus seeking and self-policing, policies (not bureaucracy but not a total anarchy either), etc.

We organize the rest of the paper as follows. We first provide a brief overview of the history and working of Wikipedia, followed by a summary of prior reviews. We then describe our literature search and article selection criteria. When summarizing the insights, we followed [17] guidelines to structure the review in three sections: (1) a summary of previous and current research, (2) an analysis of contradictions, gaps, and inconsistencies in the literature, and (3) suggestions for next steps [125].

2 WIKIPEDIA AND RESEARCH ON WIKIPEDIA

Founded on January 15, 2001, Wikipedia is a free encyclopedia written collaboratively by volunteer editors, many of whom are readers of Wikipedia as well. It is powered by the MediaWiki technology, which was designed to support online content creation with editing and hyperlinking. Wikipedia pages are organized by namespaces. A namespace is a collection of Wikipedia pages whose names begin with a particular reserved word recognized by the WikiMedia software, with the exception of article pages for which the prefix can be absent. For instance, all user pages begin with the prefix of "User". As of January 2023, Wikipedia has 30 namespaces including 14 subject namespaces (e.g., Main/Article, User, Category, and Template), 14 corresponding talk namespaces for talking about these subjects, and 2 virtual namespaces [197]. Collectively, the namespaces support knowledge creation and interactions among Wikipedia editors. Most readers browse Wikipedia pages in the "main/article" namespace. Wikipedia users can create user pages and comment on them.

Since its inception in 2001, Wikipedia went through a stage of exponential growth, with the number of active editors and the amount of editing peaked around 2007 [169], and then shifted to a slower growth or maintenance mode. As of January 2023, the English Wikipedia featured over 44 million registered editors, including over 1.1 million who had edited Wikipedia more than 10 times [187]. As the world's largest encyclopedia and the most accessible compilation of knowledge to exist in the history of humankind, Wikipedia has received both great recognition and frequent criticism since its early days. Prior research has attributed much of Wikipedia's success to a dedicated community of volunteer editors, the social structure that facilitates coordination among editors, and the technological infrastructure that makes this possible [165]. The major complaints have focused on the accuracy, completeness, and neutrality of its articles. There also have been complaints about difficulties in using the Wiki makeup editing tool, under-coverage of topics that are not of interests to its primarily young, male, and technically-oriented editor base, and the increasing bureaucracy with ever complicated rules and a hierarchy of privileged users who enforce them [205].

2.1 Prior Reviews of Wikipedia Research

Since its origin, Wikipedia has attracted interest from researchers in many disciplines. The exponential growth of Wikipedia in its early years was accompanied by a similarly exponential growth of academic interests in studying Wikipedia or using Wikipedia data for research. Google Trends [64] showed that searches for "Wikipedia research" peaked around 2006-2010, declined and remained at a lower level ever after, as shown in Figure 1.

Many reviews have tried to summarize research related to Wikipedia. The first review by Ayers [13] included 18 scholarly studies on Wikipedia and identified two main categories around content and community. Martin [117] identified key categories around article quality, trust and reliability of articles, semantic extraction, governance, and society. Bar-Ilan and Aharony [15] conducted a content analysis of over 2,000 Wikipedia studies and found that more than half of the studies were using Wikipedia articles as a data source, instead of studying Wikipedia as the central phenomenon. Similarly, Nielsen [139] identified four types of research involving Wikipedia: 1) research that examines
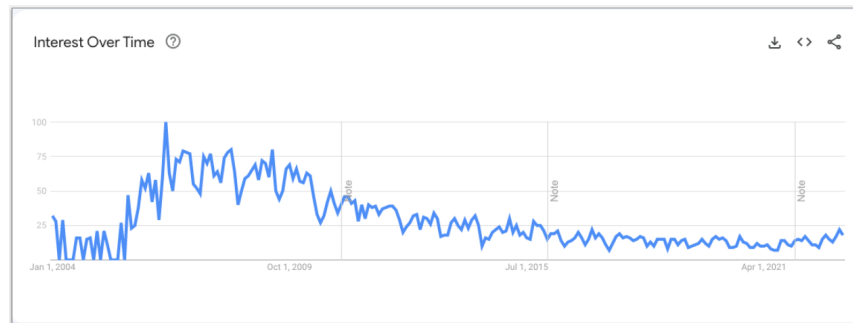
Fig. 1. Google Trends: Wikipedia Research

Wikipedia, 2) research that uses information from Wikipedia, 3) research that explores technical extensions to Wikipedia, 4) research that uses Wikipedia as a resource for communication.

Okoli et al. [141]conducted the most ambitious and comprehensive review of Wikipedia research. They reviewed 477 studies including 60 conference papers, and identified 6 topics: participation, corpus, readership, content, general, and infrastructure. Among the six topics, participation was the most popular with 200 articles examining Wikipedia contributors and their activities. The authors further identified subcategories within participation, such as motivations, cultural of collaboration, policies and governance, outcomes of participation, and software tools to support collaboration. Content was the second most popular topic, with 91 articles examining dimensions of article quality and factors that may lead to high-quality articles.

Most other reviews focused on a specific line of Wikipedia research such as community [91], content [124], readership [142], vandalism [174], or using Wikipedia as a corpus to study information retrieval and natural language processing [121, 122]. The review on content by Mesgari et al. [124] identified two themes—the quality of Wikipedia content and the size of Wikipedia— and generated two key takeaways. The first is that Wikipedia, with some exceptions (e.g., health), is comparable to many established, credible sources in terms of comprehensiveness, currency, readability, and reliability. The second is the unequal coverage across domains, with better coverage of science, music, culture, geography, and people compared to social sciences, philosophy, medicine, and law. Okoli et al. [142]'s review on readership showed Wikipedia's popularity as a knowledge source, for both light topics such as entertainment and serious topics such as health and law. They also identified concerns with the completeness and credibility of health-related articles, and noted that despite these concerns, readers generally see Wikipedia as a credible source. Vandalism refers to the addition, removal, or change of content in a deliberate attempt to damage Wikipedia. Tramullas et al. [174]'s review of vandalism showed that three quarters of the research was conducted in the computer science domain, with the primary goal of detecting vandalism using machine learning tools.

2.2 How Our Review Extends Prior Reviews

The primary focus of our review is the community of volunteer editors, i.e., their motivations, activities, interactions, and the processes through which they coordinate and collaborate. According to [132], two key factors behind Wikipedia success are "a readable product with a high degree of informational excellence" and "a thriving online community" with sophisticated social and technical mechanisms, self-governance, and a strong ideological mission. This thriving community is our focus and is similar to the participation category in [141]. Jullien [91] also conducted a comprehensive

review on the Wikipedia community by summarizing insights from over 200 journal articles around editor activities and roles, the structure and governance of the Wikipedia community, article quality, and the effectiveness of Wikipedia coordination processes. Jullien [91] also included anecdotal analyses and comparisons of the editor communities across three language editions of Wikipedia: English, French, and Spanish.

Our review extends previous reviews in at least three ways. First, it has been almost a decade since the publication of the most recent reviews in 2012, and our review includes the latest research, especially in areas that have only recently caught the attention of researchers such as algorithmic governance with bots, gender gaps in contribution, and newcomer socialization. Second, our review primarily focuses on conference proceedings, especially Association of Computing Machinery (ACM) conferences, while previous reviews focused more on journal publications. For instance, Okoli et al. [141] initially excluded all conference proceedings and only included about 60 of the highest-cited conference papers after receiving feedback from the "Wikimedia researchers' mailing list" (p. 12). Jullien [91]searched Scopus and Web of Science for studies to review, which can be limited in locating conference papers, especially those published in the ACM Digital Library [14].[2] Much of the research on Wikipedia editors' behaviors and social processes has been conducted by researchers in the field of computer science and published in conference proceedings, especially ACM conferences. Hence, we believe there is value in a review that focuses on ACM publications. Third, although previous reviews have done a good job of summarizing individual papers, our goal is to integrate across individual papers to identify higher-level insights and develop an integrative framework. We followed Bem's recommendations [17] to conduct an in-depth conceptual analysis of articles and topics using the Input-Mediator-Output-Input (IMOI) framework on team effectiveness [82] as the organizing framework.

3 LITERATURE SEARCH AND ARTICLE SELECTION

According to prior reviews, research that mentions the word Wikipedia is "extremely cross-disciplinary" [141], and much of the literature does not study the working of Wikipedia [174]. Many articles either simply mention Wikipedia in passing or use Wikipedia articles as a data source. Conducting a literature review across disciplinary boundaries is challenging [27] because different fields have different norms and outlets for publication. As mentioned earlier, we started our literature search with the ACM Digital Library to better cover works published in conference papers, especially research on the community and editor aspects of Wikipedia. We searched for articles with the keyword "Wikipedia" in paper title and abstract. Our search returned over 1,300 articles published between January 2001 and May 2020. We carefully examined article titles, abstracts, and the articles if needed and included 279 articles that focused on human behaviors around Wikipedia collaboration. We screened these articles and used snowball sampling to expand the scope of the review by including articles that appeared in the reference lists of the articles that we initially identified.

In the article search and selection process, we excluded articles: (1) that only use Wikipedia data to develop information retrieval algorithms, (2) that only introduce a tool without describing how the tool affects human behaviors, or (3) that merely mention Wikipedia without studying how it works. We also excluded most conference posters, invited talks, panels, doctoral colloquium, and workshops because they are not archival publications and might not have been peer reviewed. We included extended abstracts that reported research findings. We focused primarily on the English Wikipedia because prior research has shown significant differences between the English edition and other language editions due to differences in population and education levels, Internet penetration, and culture of volunteering [139].

---

[2][14] searched three databases: ACM Digital Library, Web of Science, and Scopus for publications on "information retrieval" between 2013 and 2016. Among the 8,699 retrieved items, more than 60% were conference papers and only 977 or 11% were located in all three databases. There was a 63% overlap between WoS and Scopus and only a 30% overlap between ACM and WoS. We counted the list of references in [91] and only about a quarter of the references were conference papers.

We believe comparison across different language editions warrants a review on its own. Our final sample includes 217 articles.

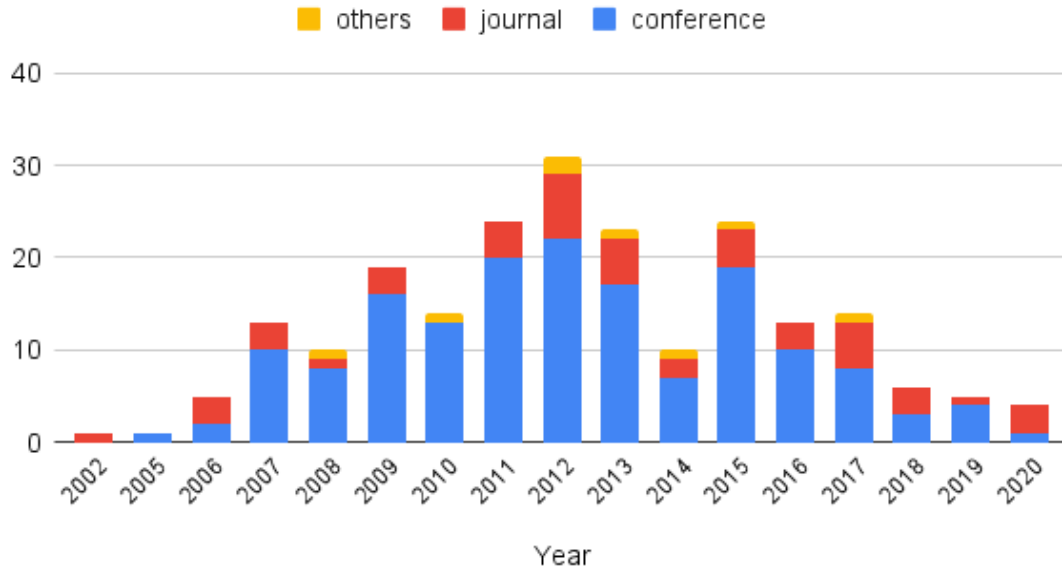## Distribution of articles across year

Fig. 2. Article Counts across Years (2001-2020)

4 AN INTEGRATIVE FRAMEWORK OF WIKIPEDIA COLLABORATION

In this section, we use the Input-Mediator-Output-Input Model on team effectiveness [82] to summarize key insights from our review. The framework represents shared understanding among team researchers about why some teams are more effective than others. Team inputs include characteristics of team members (e.g., demographics, competencies, personalities), team-level factors (e.g., team size, task structure, tools), and organizational and contextual factors (e.g., organizational design and culture). Team inputs influence the mediators such as team processes (e.g., planning, coordination, conflict resolution) and emergent states (e.g., team cognition, cohesion, trust), which affect team outputs. Team outputs include team outcomes that are valued by constituencies such as team performance (e.g., quality and quantity) and members' affective experiences (e.g., satisfaction, commitment).

The initial version of the framework was called the Input-Process-Output (IPO) framework, which was later revised to become the Input-Mediator-Output framework (IMO) and then the Input-Mediator-Output-Input framework (IMOI) [118, 120]. We chose the IMOI framework over earlier versions because (1) it considers a broader set of mediators, including both team processes and emergent states, and (2) it models not only the influence of team inputs on mediators and outputs but also the recursive feedback process of team outputs influencing subsequent inputs. The IMOI is a

well established in the organizations literature and has been used to synthesize research in other online collaboration contexts such as open source software development [36].

In this work, we combined factors in the IMOI framework and a bottom-up, qualitative analysis of article abstracts to identify the key constructs to be included in the integrative framework. In the qualitative analysis, we coded the abstracts of the 217 articles in our sample using a grounded theory approach [63]. Grounded theory is a qualitative method to analyze textual data to identify common themes and develop theory. We first conducted open coding, which involves coding the abstracts and generating categories to describe the topics of the articles. Open coding is iterative in the sense that when processing new data (e.g., new abstracts), we generate new categories and refine or aggregate existing categories to form higher-level categories [28]. In the process, we referred to factors in the IMOI framework to inform our coding. Our final set of categories include: editors, motivation, tools, conflict, coordination, governance, leadership, newcomer socialization, roles, contribution quantity and quality. We arranged these constructs using the IMOI framework by considering whether they constitute collaboration inputs, mediators such as processes and emergent states, or collaboration outputs. Working with two research assistants, we also read all articles and extracted key information such as research questions, settings, methods, participants, sample size, and key findings. Figure 3 shows our integrative framework on collaboration among Wikipedia editors. In the next section, we summarize the key insights by research topics and in the order of collaboration inputs, processes, emergent states, and outputs.
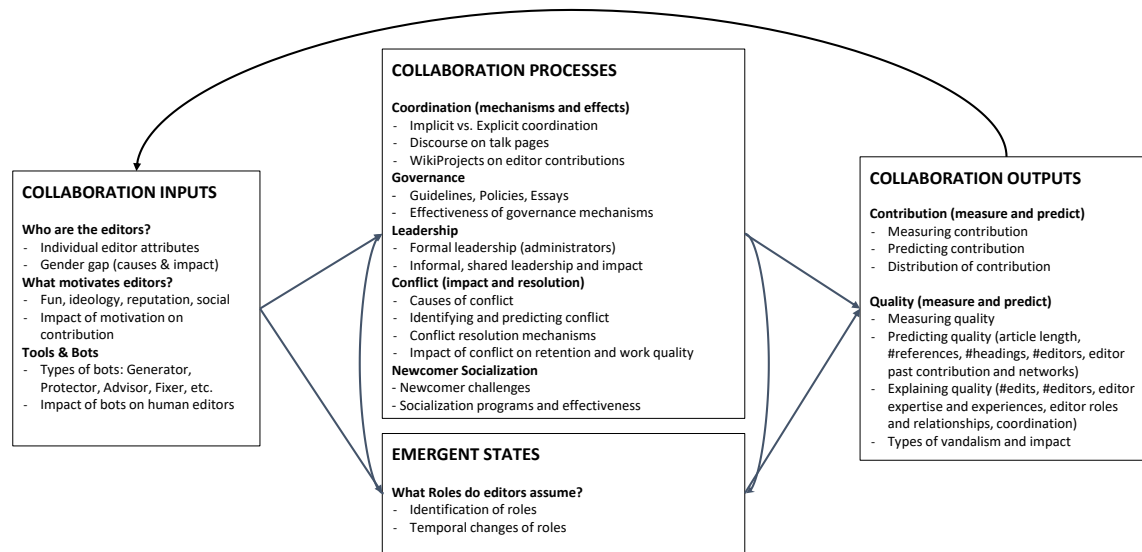

Fig. 3. An Integrative Framework of Wikipedia Collaboration

5 WIKIPEDIA COLLABORATION INPUTS

In this section, we summarize factors that serve as inputs to Wikipedia collaboration, including the editors, their motivations, group composition such as group size and diversity, and the tools used to support the collaboration.

5.1 Who Are the Editors?

Who are the volunteers who built the free online encyclopedia? By one taxonomy, Wikipedia involves three types of editors: registered editors, anonymous editors, and bots [168]. While anonymous editors constitute a large group, they contributed only a small percentage of the edits [92], such as 27.43% of labor hours or 25.83% of edits with the majority of work done by registered editors [57]. Because anonymous editors are only tracked by IP address, we know little about them. Most data on the demographics of Wikipedia editors were collected from opt-in surveys conducted by Wikimedia Foundation (WMF) and partners [186, 208]. According to these surveys, Wikipedia editors are mostly young and male, with approximately 90% male and 70% younger than 40, 50% younger than 30, and 25% younger than 21 years old. Wikipedia editors come from diverse education backgrounds with approximately one third with grade school and high school education, one third with an associate or undergraduate degree, and one third with a master or PhD degree.

Although these demographic surveys do not indicate whether editors with advanced degrees differentially contribute to articles relevant to their degree, other sources suggest that editors who contribute to specific topics often possess specialized education and expertise in that topic. For example, approximately half of the top editors contributing to medical articles were healthcare professionals or students studying in the healthcare domain [77]. The editors who have contributed to medical articles without medical training often described their work as primarily grammatical, formatting, and fighting vandalism. Wikipedia editors are also more computer savvy, spend more time online, and are heavier consumers of online information than non-editors [185].

5.1.1 The gender gap: Underrepresentation of female editors. In recent years, Wikipedia gender gap has attracted great attention. A 2011 New York Times article reported that fewer than 15 percent of Wikipedia editors were women [33]. We identified 21 articles on the gender gap, most of which published between 2011 and 2016. This research examined evidence for the gender gap, the causes behind the gap, and its impact. All studies confirmed the underrepresentation of female editors, with percentages ranging from 10 to 12% in [208] to 18% in [5]. Hill and Shaw [78]used propensity score estimation techniques to account for survey non-response bias and adjusted the female percentage from 12.1% to 16.1

Research has identified several factors that might contribute to the gender imbalance. One factor is Wikipedia's culture, which can be hostile to female editors. Compared to male editors, female editors are more likely to be reverted and more likely to stop editing [111]. They are more likely to suffer mental or emotional stress, e.g., accusations of behaving like a woman and showing too much emotion when dealing with conflict [123]. Compared to male editors, female editors were more likely to cite conflict as a reason for ceasing to contribute [34]. Despite these obstacles, many female editors continue to contribute because they value the mission of Wikipedia and experience ideology gains from helping Wikipedia and its readers [123]. Another possible reason that females contribute less than males is their perception that they lack sufficient knowledge. Although both men and women cite knowledge deficits as a reason for not contributing [100], controlling for education and experience, women are more likely to think they do not have enough knowledge to contribute [34]. Factors that have been suspected but not linked to the gender gap include the lack of Internet skills, time and availability [34, 75].

The gender gap has had important, negative consequences for Wikipedia, including but not limited to content coverage, the portrait of men and women in Wikipedia articles, and the total volume of contribution. In terms of content coverage, female-oriented topics have much less coverage than male-oriented ones [5, 111]. The gender gap has also led to stark differences in how men and women are portrayed in Wikipedia articles [179]. Specifically, the most indicative words for men often pertain to the domains in which they are active (e.g., sports or professions), whereas the most indicative words for women often refer to their gender or relationships (e.g., women, female, family). In

terms of contribution, female editors tend to perform larger revisions and have greater involvement in the social and community aspects of Wikipedia than their male counterparts [5, 111]. The low participation of female editors, hence, may compromise not only coverage and article quality but also Wikipedia's community culture.

Future research directions. Technical and social interventions to reduce the gender gap are fruitful directions for future research. Research on innovative newcomer recruitment and socialization techniques has shown some promise [46, 49, 126]. For example, Farzan et al. [46] showed that time-delimited edit-a-thons on topics relevant to women attracted newcomers and led them to continue contributing to Wikipedia after the event was over. Similarly, the WikiEd project, which encourages college instructors to use Wikipedia writing assignments in their courses, recruited a substantially higher proportion of women editors, encouraged them to make more edits, and to contribute to Wikipedia for far longer than editors who register for Wikipedia directly [114]. The Teahouse environment socializes newcomers by allowing them to create a simple profile and ask questions [126]. It attracted a higher proportion of female editors (32%) than the general editor population (16%), and increased the contribution of participants, i.e., more edits and for longer periods [131]. There is evidence that the quality gap between female versus male topics, such as the coverage of women scientists, have shrunk since 2012 shortly after media spotlight on Wikipedia's gender gap [69]. Future research should develop and evaluate programs that are specifically designed to recruit and support female editors, by e.g., improving the community culture, connecting new female editors with experienced female editors, and making editors' gender salient or invisible to fit the contexts.

Better understanding of the roles that female editors play in the collaboration process is another important future direction. Research on collective intelligence shows that having female members of a team improves team performance for many different tasks [210]. Team intelligence does not depend on the intelligence of individual members; instead, it depends on the average social sensitivity of its members and the proportion of female members. These results suggest that despite their low percentage, female editors may play a critical role in the process of article collaboration. Hence, the underrepresentation of female editors may be harming the ability of editors to collaborate.

5.2 What Motivates Editors?

Another commonly asked question about Wikipedia is why editors voluntarily contribute to Wikipedia without pay? Surveys and interviews of editors have suggested multiple motivations, including intrinsic motivations (enjoyment, flow, learning), extrinsic motivations (self-enhancement, dominance, reputation, personal use, career benefits), and social motivations (altruism, social identity, community, ideology) [16, 109, 140].

Although research has identified the common set of motivators, there are disagreements in the degree to which various motivations drive editor contribution. A common tension is between prosocial and self-interested motivations. There is often an idealized view among researchers and the public that Wikipedia contributors are primarily motivated by altruism and ideology, as illustrated by statements such as "I feel it is important to help others", "I think information should be free", "educate humanity/raise awareness" [109, 140]. However, empirical evidence shows that ideological, altruistic, and other social motivations tend not to be associated with higher levels of contribution [2, 140]. For example, in [140], editors who reported that they worked on Wikipedia for fun contributed more (r=.32. p<.001), while editors who reported being motivated by ideology (r=.11, p>.10) or altruistic values (r=.17, p > .10) did not contribute more. Instead, motivation seem to fit better with the notion of "selfish altruism" under which editors contribute for the benefit of others while also gaining psychological rewards for themselves (e.g., being heard or having fun) [3, 16, 140].

Empirical evidence from field experiments provides further support that contribution to Wikipedia is often driven by self-interest and extrinsic motivations such as recognition. For instance, economists who were invited to contribute to

articles that matched their expertise were 13% more likely to contribute when they were informed that doing so would help increase citations of their work [29]. Wikipedia editors' contributions have also been linked to general propensity of reciprocity and preference for social image, not altruism [2]). Symbolic recognition, such as "barnstars" (i.e., the Wikipedia awards that one editor bestows on another) have been shown to substantially increase editor retention and contribution, and the positive effects persisted over four quarters [54, 155].

Editors' motivations change and evolve over time, as editors gain experiences and become more identified with Wikipedia. Compared to newcomers who often see Wikipedia as a collection of articles, experienced editors tend to see Wikipedia as a community of editors [21] and are more likely to report social and community factors as important motivations [16]. Compared to general editors, Wikipedia administrators often place greater value on social motivations and the creation of Wikipedia articles as public artifacts [16]. Interestingly, recognition and social image seem to be important drivers for both new and experienced editors. New editors who value reputation tend to take on more roles and get involved in more articles than editors who value other types of motivations [7]. Preferences for social image has been associated with increased contributions for experienced and highly engaged editors [2].

Finally, the current literature also suggests several reasons for the lack of contribution. Although some editors may want or be willingness to help, they may not contribute due to the lack of time, lack of qualification, and unfamiliarity with the markup language [109].

Future research directions. Future research may focus on understanding the dynamic evolution of editors' motivations both within their life cycle and within the life cycle of an article or the platform as a whole. Most of the research we reviewed examined editors' motivations in the early or peak days of Wikipedia, i.e., 2007 and prior. What motivates editors now may differ from what motivated them then. Contextual factors such as article topics or types of tasks may also influence motivations. What motivates an editor to contribute to an already well-developed article may differ from what motivates them to contribute to an undeveloped, stub article. It is also important to examine different types of contribution besides editing articles, such as coordinating work on talk pages, welcoming newcomers, vetting new articles for their quality, or resolving disputes. Future research should try to unpack the heterogeneity among editor motivations and their effects on contribution of different types.

5.3 Bots and Their impact on Wikipedia Collaboration

Wikipedia bots are autonomous software programs that are developed and operated by human editors to perform tasks around article editing, maintenance, and administration [60, 216]. While the number of human edits have been decreasing from 2006 to 2012, bot edits have been steadily increasing, with 40% of bot edits occurring in non-article namespaces [134]. The first notable bot on Wikipedia, called RamBot, appeared in 2002 to create 30K city articles based on U.S. census data [73]. As of February 2019, Zheng et al. [216] identified 1,601 registered bot accounts, many of which had made thousands to over a million edits. Most bots perform repetitive tasks like fixing links or grammar errors, capitalizing unique proper nouns, sending notifications or messages, and fighting vandalism. Over the years, edits by bots grew from 4-5% in 2006 to 16.33% of all Wikipedia edits [55]. For example, ClueBot NG is an anti-vandalism bot that detects and reverts vandalism quickly and automatically. It has been operational since 2010, made millions of edits, and is the 6th most prolific editor in the English Wikipedia [58].

There are established procedures to introduce bots on Wikipedia. The volunteers who develop a bot need to first submit a bot approval request to the Bot Approvals Group consisting of experienced editors, which vets new bot proposals and addresses bot-related grievances. The request needs to include information about the bot's functions, programming languages, and estimated number of pages affected. If the group decides that the bot is helpful and follows

Wikipedia policies, the bot will be approved for a short trial period during which it will be closely monitored and if there are no issues, it will be fully deployed afterward.

5.3.1   Types of bots. There have been several taxonomies to classify bots. One taxonomy distinguishes between fully automated bots and semi-automated bots, and the latter also known as assisted editing programs [60]. The combination of human editors and assistance tools working together is referred to as "tool-assisted cyborgs" [60]. ClueBot NG is a fully automated bot, built on Bayesian neural networks to automatically detect and revert vandalism. It is powerful enough to scan every edit made to Wikipedia in real time [58]. In comparison, Huggle is an assisted editing program and presents human editors with a list of edits ranked by their likelihood of being vandalism. The human editors can review the edits and revert them with a single click, and they can automatically leave a warning message on the offending editor's user talk pages.

Another taxonomy classifies bots by the tasks they perform. Halfaker and Riedl [73] identified four sets of activities including injecting public domain data, monitoring and curating content, augmenting the MediaWiki software, and protecting the encyclopedia from malicious activity. Zheng et al. [216] developed a comprehensive bot taxonomy using unsupervised machine learning methods and identified nine categories of bot functions:

- Generator (generate redirect pages or article pages based on other sources)
- Fixer (fix links, content, files on article pages)
- Connector (connect Wikipedia with other sites and databases)
- Tagger (patrol articles and tag articles with templates and categories)
- Clerk (update statistical information or maintenance pages, deliver article alerts to Wikiprojects)
- Archiver (archive closed discussions)
- Protector (detect and regulate destructive behaviors like spam and vandalism)
- Advisor (provide suggestions to editors and Wikiprojects about articles to improve)
- Notifier (deliver messages to editors)

Zheng et al. [216] found that based on the number of bots and edits by a particular type of bots, Fixer is the most common type of bot, followed by Tagger and Clerk, and then Connector, Notifier, Protector and Archiver. Generator and the Advisor are the least common. On the other hand, Protector is the best known and most studied bot type.

5.3.2   Impact of bots on human editors. Research suggests bots have both positive and negative effects on human editors, with most of the research focused on the impact on new editors' survival or their contributions [216]. Zheng et al. [216] analyzed three types of bots that directly interact with human editors and found that Protector bots generally had negative effects on newcomer survival, Advisor bots had positive effects, and Notifier bots had no effects.

There is clear evidence for the negative effects of Protector bots and algorithmic vandal fighting tools on newcomer retention [70, 216]. Halfaker et al. [70] attribute the reduction in the number of active editors on Wikipedia starting about 2007 to the growth of algorithm tools especially vandal fighting bots. Reverts made by a bot have a substantially larger effect of driving away desirable new editors than reverts made by a human editor [70]. Meanwhile, not all Protector bots had negative effects on newcomer retention. Compared to ClueBot NG, XlinkBot, a bot designed to automatically revert the addition of external links that are spam or do not comply with Wikipedia policy, did not have any negative effects [216]. It was possibly because the messages it sends to editors were longer, more friendly, and informative (e.g., contained the specific reason for the reversion, links to the guidelines, information about how the bot works and the bot creator's FAQ page).

Despite the negative effects of ClueBot NG on newcomer retention, it plays an important role in patrolling and protecting Wikipedia as a source of high-quality information. Its value became especially visible when the bot went down for four distinct periods in 2011. Geiger and Halfaker [58] treated these breakdowns as natural experiments and found that overall time-to-revert edits almost doubled when the bot was down. Although most problematic edits were eventually reverted, suggesting that other people or programs took over the quality control work, the reverting happened at a far slower rate.[3] These results suggest a critical trade-off between the efficiency of vandal fighting and the costs of discouraging newcomers.

In contrast to Protector bots, Advisor bots generally have positive effects. For example, SuggestBot, an Advisor bot, used text analysis, collaborative filtering, and hyperlinks to recommend articles for editors to contribute to. These targeted recommendations increased the number of edits by roughly four times compared to suggesting random articles [35]. HostBot, another Advisor bot, increased the survival rates of newly registered editors by inviting them to join the Wikipedia Teahouse [131]. LeadWise, a bot designed to recruit experts to contribute to Wikipedia in their areas of expertise successfully mobilized experts to contribute and interact with one another [49].

Wikipedia to some extent has become an ecosystem of humans and bots working together [216] or a decentralized network with humans and bots acting in coordination with each other [60]. Using trace ethnography, Geiger and Ribes [60] described a case where four human editors and the Huggle bot acted together to identify and ban a vandal in 15 minutes. They concluded that bots and assisted editing programs are significant social actors in Wikipedia, making possible a form of distributed cognition. Huggle's automated actions, such as issuing a stronger warning after detecting a previous warning or recommending the banning of a vandal to human editors, were crucial to the functioning of the distributed system.

There have been concerns with unintended consequences of human-bot or bot-bot interactions on Wikipedia, which may cause bot wars in which bots repeatedly revert one another. Tsvetkova et al. [175] found over a ten-year period, bots on English Wikipedia reverted another bot on average 105 times, significantly higher than the average of 3 times for human editors. Because bots are more active than humans, the higher number does not mean that bots fight more than humans. In fact, compared to humans, a smaller proportion of bots' edits were reverts. Geiger and Halfaker [59] replicated the study and showed that bot-bot reverts were only 0.5-1% of all bot edits and most bot-bot reverts were not reciprocated. Only 1.46% of all bot-bot reverts to English Wikipedia were cases where two bots reverted each other more than twice in a single article, which might have constituted bot-bot conflict. Even these cases were the outcomes of unintentional designs by the bot developers to work on opposing tasks. Once uncovered, the developers were able to resolve the conflict via talk page discussions.

Future research directions. We have much more to learn about Wikipedia bots. One future direction is to study how humans perceive and respond to bots. There has been evidence that Wikipedia new editors often do not know if they are interacting with a human or a bot [50]. How common is it for human editors to be aware of bots' presence and prevalence on Wikipedia? How do human editors perceive bots? What happens when bots make mistakes and how do human editors respond to human errors and bot errors similarly or differently? Another interesting direction is to study Wikipedia as a complex adaptive system, in which humans and bots learn from and adapt to one another [216]. Existing research has only scratched the surface on some first-order effects that bots have on human editors, and there may be second-order effects on human-human interactions, which may in turn affect human-bot interactions.

---

[3]The study also compared the time-to-revert for manual reverts (1 minute to 24 hours), assisted reverts by Huggle (roughly within a minute), and automatic reverts by ClueBot NG (mostly within 30 second and the majority within 5 seconds)[58].

Another important future direction is algorithmic bias. There has been evidence showing that some bots have algorithmic bias against anonymous or new editors [60]. Users whose edits have been previously reverted are viewed by the algorithms as even more suspicious than generic new editors; editors who have been left warnings on their user talk page (a process explained below) are automatically sent to the top of the queue of suspected vandalism. How fair or effective are these practices? What are the long-term effects on Wikipedia? Recent research has proposed value-sensitive algorithm design, which aims to balance the goals of reducing community maintenance work, maintaining human judgement as the final authority, supporting different workflows, encouraging positive engagement with diverse editor groups, and establishing trust between human and algorithms within the Wikipedia community [167, 219].

6 WIKIPEDIA COLLABORATION PROCESS

This is the most central part of our review, which focuses on coordination and collaboration among editors. Coordination defined broadly means managing interdependencies among activities [116]. There is a high degree of interdependence among the activities of Wikipedia editors, especially those working on the same articles. We organize this section around coordination mechanisms, governance policies, leadership behaviors, conflict detection and resolution, and editor roles. Although we cover each topic in a subsection, the topics are highly intertwined and understanding the whole of Wikipedia collaboration requires considering all the topics together. The subsection on coordination focuses on a narrow aspect of coordination - talk pages and WikiProjects - as two common coordination mechanisms. The governance subsection focuses on policies, guidelines, and collective decision-making. The leadership subsection focuses on formal leadership roles like administrators and shared leadership. The roles subsection focuses on the prototypical behavioral patterns that define less-formally defined roles. Finally, despite the various coordination mechanisms and policies, conflict still occurs, and the last subsection focuses on the detection and resolution of conflict.

6.1 How Do Editors Coordinate Their Work?

As Wikipedia grew in size and complexity, coordination work in Wikipedia grew at a pace faster than the growth of articles [177]. We identified 29 articles related to coordination, most of which focused on communication among editors on article talk pages and WikiProject discussions [103, 177]. Some articles mentioned complementary venues, such as Internet Relay Chat (IRC) channels, Village pump, bulletin board, comment pages, mailing lists [148]. However, these channels were rarely the focus of a study because most are ephemeral and do not retain activity logs.[4] Article talk pages are often used to facilitate and assure article quality whereas WikiProject discussions are often used to encourage editor contribution to articles related to the project. Hence, we review the two separately.[5]

6.1.1 Article talk page discussion and its impact on article quality. Most articles, especially those with more editors and editing activities, have active talk pages [160]. On the scale of Wikipedia as a whole, editing on article pages and commenting on talk pages have co-evolved "nearly in perfect synchrony" with similar activity cycles ([93], p. 3). In other words, activities across all article pages on Wikipedia tend to fluctuate with activities on article talk pages. However, at the individual article level, article editing and talk page commenting tend to be independent of each other. Peaks in the editing of an article do not necessarily correspond to peaks in commenting on the article talk page.

---

[4]Some channels, however, like IRC played an important role in coordinating work around breaking news articles because they facilitate synchronous real-time discussion and can serve as a centralized command and control center [97].
[5]Researchers have also examined coordination across languages [67, 76], in particular, how editors contribute to articles on similar topics across different languages [67] and how to properly translate articles across languages [76]. We did not include these papers because of our primary focus is on the English version of Wikipedia.

Content analyses of article talk pages show a variety of themes, such as requests and suggestions for editing coordination, requests for information, pointers to Wikipedia guidelines and policies, references to internal and external sources, polls on controversial edits, requests for peer review, and off-topic remarks [177]. The most common use is requests and suggestions for editing coordination, accounting for 58.8 to 72% of the comments [45, 160, 177]. Editors post to the talk page to discuss their editing plans, ask for help, or explain the reasons why they think specific changes should or should not be made. Next in frequency is requests for information (10.2%) and references to Wikipedia guidelines and policies (7.9%) [177]. Requests for information are generally from editors who have no intention to edit the article but post on the talk page to tap into expert knowledge on specific topics (e.g., asking about the role of political parties in a UK election for a class report).

How does talk page discussion and coordination affect article quality? It depends on the type of coordination. Kittur and Kraut [102] examined two types of coordination. Explicit coordination involves discussion among editors and is measured as the number of edits made to article talk pages. Implicit coordination occurs when a small number of editors set directions by making a disproportionate number of edits to an article and is measured by the Gini coefficient of the concentration of edits across editors. Both explicit and implicit coordination led to greater quality improvement, with the associations stronger early in an article's life history.[6] However, as the number of editors increases, the benefits of explicit communication began to decline whereas the benefits of implicit coordination continued. In other words, with more editors to an article, article quality was higher when work was highly concentrated than when work was evenly distributed across editors [102]. They also found that over five times as many editors edited the article than posted in the article talk pages, meaning only a small subset of editors participated in article discussions.

6.1.2 *WikiProject coordination and its impact on editor contribution.*

Another coordination mechanism is the WikiProject, which is a social mechanism to mobilize editors and facilitate collaboration around articles on certain topics or around certain tasks. Some WikiProjects are topical (e.g., arts, music, Africa, biology, technology) and others are task-oriented (e.g., fixing typos, decisions about new articles, social and community support). Gilbert et al. [61] described WikiProjects as a way to lower coordination costs of monitoring and task routing by providing a centralized repository of tasks and discussion for explicit group coordination. Besides coordination, WikiProjects also provide opportunities for editors to find and socialize with other editors and feel a sense of belonging to a smaller community within Wikipedia [51].

Coordination at the project level is more challenging than coordination at the article level, because project coordination involves tracking and managing hundreds or thousands of articles over time. Since 2007, editor participation in non-topical projects has remained stable where editor participation in topical projects has been decreasing [129]. An important goal of a WikiProject is to mobilize editors and increase their contribution to articles associated with the project. Similar to implicit coordination in article creation [102], WikiProjects were more effective in getting more editors to contribute to the project when a small number of project members coordinate and structure the project [51]. After joining a WikiProject, editors were more likely to work on project-related content, to shift their contributions toward coordination rather than just article editing, and to perform maintenance work such as fighting vandalism [103]. Members who joined a project also increased their interactions with other project members, i.e., edits on each other's user talk pages, and the effects lasted five months after joining a project [51].

Content analyses of WikiProject talk pages showed a somewhat different pattern from analyses of article talk pages. Only 7% of the discussions contained explicit requests to work on the project whereas most messages were requests

---

[6]In a follow-up study, [8] found that local inequality in contribution to articles increased coordination via talk pages but had no direct effect on article quality, which contradicted [102]. We believe the differences are due to different sampling strategies with [102] sampling 147,360 articles across all quality levels and [8] sampling 50 articles on science and technology

for opinions (38%), request to join discussion to resolve disputes (16%), or to make the group aware of one's activities (18%) [130]. Interestingly, non-members participated actively in WikiProject discussions and initiated more than half of the discussion. Members responded to non-members in roughly comparable ways [128]. Another WikiProject coordination mechanism is Hot Articles links, which shows a ranked list of articles that have had frequent editing activities. Comparing the two, Gilbert et al. [61] found that being included on project talk pages increased the number of unique editors to an article and being included under Hot Articles links increased the number of edits.

Future research directions. Our review shows both the importance of coordination and the need for future research to understand how to improve coordination among editors. While talk pages and WikiProjects allow editors to be aware of others' editing activities and identify needed work, more advanced tools are needed to facilitate complex interactions among editors, such as better awareness and management of task dependencies [62]. Another fruitful future direction is to understand the social dynamics on talk pages affect editors, especially newcomers. Most existing research has focused on classifying content of discussions on talk pages, with little assessment of the sentiment, tone, norms, social dynamics, and the network structure of the talk page discussions, how editors make sense of these discussions, and how talk page discussions affect editors' motivations to contribute. Anecdotal evidence suggests that new editors often find talk pages confusing to navigate, but find them useful in learning about articles, Wikipedia policies, and terminology once they figure out how to navigate them [160].

6.2 What are the Governance Mechanisms among Editors?

Governance generally refers to "a system for organizing the rules that regulate people's behaviors in a particular place" ([52], p. 51). It can also refer to "the creation of order to achieve mutual gains in potentially conflict-laden contexts" ([134], p. 81). Although Wikipedia started as a self-organizing meritocracy, over time, it has established its own "bureaucratic" norms and policies that govern the behaviors of and interactions among editors [26].

Three governance mechanisms have emerged in Wikipedia: policies, guidelines, and essays [127]. Policies and guidelines are more formalized than essays, because (1) they are developed through a process of open discussion, polling, and collaborative editing with community consensus, (2) they carry official weight and violations of a policy or a guideline can have consequences, e.g., editors being banned from the website. Policies are even more formal than guidelines and allow fewer exceptions. Compared to policies, guidelines are more open to debate and likely to have occasional exceptions [108]. Essays are "opinion or advice" of individual editors or editor groups, which may not be accepted by the community.

6.2.1 Guidelines, policies, and essays. Most empirical research has focused on policies and guidelines. As formal governance mechanisms, policies and guidelines address issues such as article contents, editor behaviors, dispute processes, and what constitutes legitimate sources [90]. They serve many functions such as facilitation of editor collaboration, construction of meaning and identity, internal and external signals, and control [26]. By institutionalizing negotiated settlements and best practices as policies and guidelines, editors can eliminate future need for direct communication or reduce the costs of the communication. Instead of discussing and resolving every disagreement as an ad hoc decision, editors can simply cite policies and guidelines to persuade others [127]. The most-cited policies and guidelines are about critical issues such as citing sources, achieving consensus, preventing bias, and regulating editors' behaviors [108].

As socially constructed entities, policies and guidelines are dynamic in nature and reflect the evolving principles to facilitate collaboration among editors. Editors create, debate, and revise the policy and guideline pages as much

as they do with article pages [108]. As Wikipedia grew, both the number and complexity of policies and guidelines increased significantly. Many policies and guidelines were developed to respond to either external complaints (e.g., copyright guidelines in response to non-permitted use of protected materials) or internal conflicts (e.g., the rule that an editor must not perform more than three reverts on a single page in 24 hours in response to edit wars). Due to their dynamic and evolving nature, the counts of policies and guidelines are constantly in flux with different sources supplying different numbers. For instance, Butler et al. [26] counted 44 policy pages and 248 guideline pages whereas Kriplean et al. [108] identified 42 policies and 24 guidelines. A few years later, Joyce et al. [90] reported 52 policy pages and 147 guideline pages; Müller-Birn et al. [134] reported 383 policy pages and 449 guideline pages; and Morgan and Zachry [133] reported 511 policy pages. These discrepancies occurred partially because of the different criteria for defining a policy (e.g., "policies" versus "policy pages" with some policies spanning across multiple pages) and partially because policies and guidelines are constantly changing.

A robust temporal trend is the increasing prevalence and importance of essays [79, 133]. Unlike policies and guidelines, essays are less formal and do not require consensus among editors. Wikipedia editors write essays for a variety of reasons: to share advice or opinions, to describe an interpretation of a policy, or to assert an ideological stance. Essays often use humor, hyperbole, and anecdotes to convey serious messages about proper editor behaviors and best practices [127]. For example, Morgan et al. [127] described an essay titled "No Angry Mastodons," which conveys the editor's interpretation of the Civility policy and advises editors to avoid editing when stressed, hungry, tired, or drunk. The recent growth of essays may be the result of fewer opportunities for editors to contribute to policy creation and essays being a softer regulating mechanism, which is less formal, heavy-handed, or intimating.

Policies, guidelines, and essays are rule-based governance mechanisms. Other governance mechanisms focus on processes and reinforcements of the rules such as Article for Deletion (AfD) and page protection. In general, administrators and experienced editors are more involved in governance work than are new editors. For example, the majority of participants in article deletion debates were long-time editors [56]. Administrators have privileges and can take actions to protect pages or ban vandals. Only administrators can apply page protection. A full protection restricts the page so that it can only be edited by administrators whereas semi-protection restricts the page so that it can be edited by any confirmed editors (i.e., accounts with more than 10 edits that have existed for more than 4 days). The duration of the protection can vary from seconds to up to 6 years with a median of 14 days. Only a tiny portion of Wikipedia articles (e.g., 0.36% of English articles) have gone through page protection, including some that had undergone multiple protections. Protections have been applied disproportionately to articles with heavy readership, with some protected pages are among the most viewed pages [79].

6.2.2 Effectiveness of governance mechanisms. How effective are these governance mechanisms? The answer is mixed. On one hand, policies and guidelines have played an important role in the daily operations and overall success of Wikipedia [26]. For example, the institution of the three-revert rule reduced reverts in half, by limiting the number of times editors can revert each other's changes [23]. On the other hand, Kriplean et al. [108] found numerous instances of ambiguous policies causing different interpretations (e.g., the Neutral Point of View policy and how to write neutrally about a polarizing topic) and complex power plays, in which individuals or groups tried to claim legitimate control over an article by imposing their interpretation of policy on it. The power plays might revolve around article scope, legitimacy of contributors, or legitimacy of sources. For example, in an argument about article scope (e.g., whether views of the Catholic church belong in an article on a scientific theory), a power play might emerge from ambiguities

and inconsistencies among three different policies on neutral point of view (NPOV) [198], boldness in writing Wikipedia [192], and the policy "wiki is not a paper" [189].

Due to the complexity of Wikipedia policies, editors have also created a policy called "ignore all rules" or IAR [188], which states that "If a rule prevents you from improving or maintaining Wikipedia, ignore it" [90]. In general, IAR has functioned as a workaround and positive mechanism to help editors manage the complex interplay among Wikipedia policies, especially after Wikipedia has grown to be large and complex.

As Wikipedia grew larger, many governance functions have been delegated to intelligent agents or bots, known as algorithmic governance, in contrast with social governance by human editors [134]. for example, German Wikipedia uses bots to automatically approve and grant privileges to human editors [134]. Once an editor meets all criteria (e.g., 150 edits or 8 different article pages), they will automatically be granted the reviewer status and associated privileges.

Comparing the two mechanisms – social versus algorithmic methods for maintaining the "moral order" of Wikipedia – algorithmic governance with bots is easier to enforce and can quickly scale up whereas social governance with human editors can handle exceptions better [60]. Software algorithms may suffer from algorithmic bias, such as automatic decision criteria that discriminate against anonymous or newly-registered editors [60], and, as discussed in Section 5.4.2, some of these the algorithmic tools used to reject contributions have been linked to decline in active contributors [70].

Future research directions. Existing research on governance is quite comprehensive and thorough. Future work could examine the dynamic evolution of individual policies and guidelines as well as the overall policy or guidelines space. What triggers a new policy? How does discussion around a policy evolve and affect the development of the policy pages? Future work could also investigate how newcomers learn about policies and guidelines and develop interventions to help newcomers navigate the sophisticated policy space. More research is also needed to understand essays - what drives their creation, how are they used, and their impact on collaborative outcomes.

6.3 How Does Leadership Manifest in Wikipedia Collaboration?

Leadership in general has been defined as influencing members of a group towards a goal [22]. Leadership in online open collaborations like Wikipedia and open source software development can take multiple forms and is not necessarily associated with formal leadership roles. Merit plays an important role in the emergence of leaders. Instead of vertical authority over people, leadership is often associated with lateral authority over tasks and activities [38]. We identified 13 articles on leadership. Some of these focused on formal leadership roles, primarily administrators, and the factors important to becoming an administrator. Others focused on informal or shared leadership, and how non-administrator editors engage in leadership behaviors and influence others.

6.3.1 Formal leadership and Request for Administrators (RfA). Administrators are the main group of formal leaders in Wikipedia. They can block user accounts and IP addresses from editing, protect pages from editing, edit fully protected pages, delete pages, rename pages without restrictions, and use certain tools. To become an administrator, an editor must submit a request, undergo community scrutiny, and obtain sufficient community consensus. Research on formal leadership has focused largely on the role of administrators.

The Request for Administrator (RfA) process is the route to become a Wikipedia administrator [201]. It is an open voting process, during which any editor can view the nominated editor's record, question the nominee, and evaluate his or her ability to appropriately use the administrator rights. A successful RfA requires a clear consensus that the nominated editor "is committed to Wikipedia and can be trusted to know and uphold its policies and guidelines" [196]. The criteria include breadth and duration of editing experience, ability to work with other editors, and understanding

of Wikipedia norms and policies. The RfA process consists of three steps: a nomination statement, Q&A about past and future behaviors, and statements of support, opposition, or neutrality by community members [24]. Although there is no objective standard for the percentage of supporting votes required to get promoted, history has shown that a minimum of 80% supportive votes is typically needed [150].

Research on RfA has shown three factors key to a successful RfA. The first is strong editing experience. Among nominated editors, those who were promoted had roughly twice as many edits as those who were not promoted (3038 vs. 1604), and every additional 3804 edits increased the chance of promotion by 10% [24]. Fewer than 50% of nominees with fewer than 2,500 revisions became administrators compared to 70% of those with at least 6,000 revisions [150]. The second factor is diverse editing experiences, that is, edits to different namespaces such as article, user talk, and deletion discussion. Edits in each additional namespace increased one's chance of promotion by 2.8% [24]. The third factor is edit summaries. Editors who frequently summarized their edits and left coordination notes for future editors were more likely to be promoted [24, 39]. A few other factors also slightly increased the chance of promotion, such as longer tenure and politeness in talk page discussion (e.g., the use of "thanking" words) [24, 39].

Interestingly, some factors emphasized in Wikipedia's guidelines for administrators had no effects or even negative effects on RfA success, such as reverting vandalism, alerting administrators about page protection, or voting on article deletion [24]. Contributions to dispute resolution venues such as Arbitration or Mediation Committee pages and the Administrators' Noticeboard decreased one's chance of promotion [24]. The most uncertain factor is social interaction or edits on users' talk pages with some studies showing a positive effect [150] and others showing no or a negative effect [24, 39]. The positive effect in [150] was largely due to interactions with existing administrators on their user pages, instead of ordinary editors.

Both the success rates of RfA and the criteria have changed over time. From 2006 to 2008, success rate dropped from 75.5% to about 44.7% [150]. After 2006, contributions to WikiProject policies and non-article pages became significant predictors of RfA success, indicating a shift to prioritize policy-making over simple article contributions [24, 25]. The impact of an additional 1,000 article edits on the probability of promotion dropped from 15.4% to 6.3% from pre-2006 to post-2006, while the impact of a 1000 edits to Wikipedia policy pages grew from .4% to 19.6% during the same time period [24, 25].

Compared to research on predicting the success of promotion requests, less research has studied the impact of becoming an administrator on an editor's subsequent behaviors. Once they are promoted, do administrators contribute more or less than they had in the past? Do they change the types of contribution they make? One study suggested that after being elected, administrators focused on more controversial topics than before their promotion and seemed to be pushing positions that violated Wikipedia's policy on neutral points of view [39]. For example, administrators who focused on articles about US politics inserted partisan phrases that are "either Democratic or Republican talking points." Although these controversial edits were a small set among their total contributions, these edits might have shaped information on many sensitive topics on Wikipedia. Neither the editors' prior editing history nor the percentage of positive votes they received in the RfA process predicted which administrators were likely to demonstrate such behavioral changes. One promising way to identify suspicious candidates is to give more weights to influential voters in the voting process possibly because these voters were more informed and had gathered extra information about the nominated editors than the rest [39].

6.3.2 Informal and shared leadership. Besides formal leadership, leadership in online open collaboration also takes the form of shared leadership, when editors regardless of their formal status engage in "a dynamic, interactive influence

process" with the objective of achieving collective goals [218]. In contrast to vertical leadership by designated leaders, under shared leadership, the agents of influence are often peers who exhibit leadership behaviors such as directing, communicating, encouraging, giving feedback, and reprimanding [147]. Shared leadership on Wikipedia involves similar behaviors. For example, Zhu et al. [218] coded messages Wikipedia editors left on each other's user pages and identified four shared leadership behaviors: (1) directing (instructions, commands, assigning tasks, and setting goals), (2) providing positive feedback (acknowledging work and providing rewards), (3) providing negative feedback (regulating people through reprimands), and (4) social exchanges (social conversations and friendly talk). Both administrators and non-admins send these messages, with non-admins sending more than admins (60% vs. 40%). At the individual level, however, average admins sent over 32 times more leadership messages that average non-admins. The majority of the messages were task-oriented, over person-oriented ones (79.9% versus 20.1%) [218].

These leadership messages, whether performed by admins or non-admins, significantly affected the contributions of the editors who received them. Compared with editors who received no messages, editors who received messages with positive feedback, directing, and social exchanges increased their edits by respectively 19%, 23%, and 8.6%; and editors who received negative feedback reduced contributions by 14% [217]. The leadership behaviors of admins were more influential than non-admins, both positively and negatively [217]. Receiving a positive feedback message or a social exchange message from an admin increased editing an additional 2.8% and 5.8% over similar messages from a non-admin. Receiving a negative feedback message from an admin decreased editing an additionally 9.5% over similar messages from a non-admin.

Future research directions. While existing research has focused on the promotion to administrator status, we need research to understand the demotion process, i.e., how and why editors may lose administrator privileges. Did they give them up voluntarily or have them stripped away? As of now, we have very limited knowledge[200]. We also need research to understand what happens to administrators after their promotion. Do they increase their quantity and quality of contribution? Do they change the types of contribution, e.g., becoming more involved in community building and social exchanges with other editors? Because leaders interact with and influence others primarily through messages, it would be valuable to perform more in-depth analyses of the nature of the communication that is most effective at influencing others. For example, how does politeness, assertiveness, or appeals to common interests affect the communication of positive and negative feedback or directive messages? Do formal and informal leaders use the same techniques and are they equally effective?

6.4 How Do Editors Resolve Conflict?

Conflict is an extensively studied topic, and we identified 34 articles that examined the causes and impact of conflict, the mechanisms for resolving conflict, and the measurement and prediction of conflict or controversial articles.

6.4.1 Causes and manifestations of conflict. Conflict on Wikipedia may occur due to contentious article topics, editors' aggressive behaviors, or both [178]. Some topics are inherently polarizing such as abortion, global warming, homosexuality, and biographies of living people [89, 163]. Not all articles related to controversial topics show high levels of conflict, and the same topic may be controversial in one language but not in another [84]. When collaborating on the same article, editors need to decide what text, links, and images to include, how to structure the page, and whether the content represents a neutral point of view. Conflict can arise if editors have different views on the content, source, structure, and wording of articles or have different interpretations of Wikipedia policies. Editors with different levels of

experience or different types of expertise (e.g., domain expertise vs. expertise with Wikipedia policies) may disagree and engage in public arguments [53].

Territoriality is another trigger of conflict on Wikipedia [71, 172]. Editors can use the "Maintained" template to indicate their active contributor status and communicate their commitment to an article. Territoriality may be beneficial during labor-intensive periods of article creation and improvements, but it can increase the chance of new editors being reverted and discouraged from future contribution. The negative effects happen when the territorial editors engage in defensive actions, e.g., monitoring article watch lists for changes and vetting unknown editors.

Conflict can manifest in multiple forms: heated debates on talk pages, reverts among editors, or edit wars when mutual reverts escalate out of control [212]. Several factors influence the likelihood of being reverted, such as the quality of the contributions, experience of the contributors, and ownership of article content. Reverts are more likely to happen to edits that violated Wikipedia policies, edits that contained errors in spelling and syntax [158], edits that removed words with high word persistence (i.e., words that have survived more revisions), edits that removed words by active editors, and edits by editors who had been reverted recently [71]. In contrast, editors' active contribution experience, measured as the number of days they had been active and the number of editing sessions they had completed, reduced their likelihood of being reverted although the number of days since an editor began editing did not [71].

6.4.2 Identifying and predicting controversial articles. To warn editors about controversial articles, Wikipedia editors and administrators can manually tag articles as controversial or having disputed content. With recent advancements in machine learning, researchers have developed machine learning algorithms to predict levels of controversy, often using the manually applied controversial tags as the ground truth to train and evaluate the machine learning models [104, 212].

Machine learning models that combine multiple indicators can be effective in identifying controversy. The factors that have been linked to article controversy range from the revision history of article pages and talk pages, to editors' behaviors and interactions, and collaboration networks among editors (e.g., [104, 163, 178]). The simple measures often include the number of edits, revisions, unique editors, anonymous edits on article pages and talk pages [104]. More sophisticated measures include the ratio of deletions to edits, counts of mutual reverts among editors [170], article-level bipolarity networks showing two groups of editors with opposing views [20]), and collaboration networks showing editors' positive and negative attitudes toward one another [163].

Several patterns are evident across studies. First, the revision history of talk pages is more informative than the revision history of articles pages in predicting controversy. For example, Kittur et al. [104] found that the top factors predicting article controversy were associated with talk pages, including the number of revisions, the number of minor edits, and the number of unique editors on the talk pages.

Second, editors' interactions and networks, especially negative interactions such as mutual reverts, were more informative than simple counts of revisions [170]. Models based on editor networks showed significant improvements over models based on statistics of page revisions. Models of the number of deleted words between editors outperformed models of simple counts of edits or editors [178]. Different methods have been used to construct editor networks. Brandes et al. [20] used deletion of edits to determine negative links between editors and found that controversial articles often had a network of two clusters of editors with negative views of one another. Sepehri-Rad and Barbosa [163] used the votes that editors cast in Wikipedia administrator elections to determine positive and negative connections with a "support" vote indicating agreement and an "oppose" vote indicating disagreement. A classifier that combined

multiple measures of collaboration networks (e.g., # of nodes, # of positive edges, % nodes with high degrees, and triad features) was the most accurate in predicting article controversy.

Third, controversy is too complex to be measured by a single factor, and the most accurate models are meta-classifiers that combine many factors. The term meta-classified was coined by Sepehri-Rad and Barbosa [163] although the idea was first explored by Kittur et al. [104] who used "a combination of page metrics" to predict article controversy. Comparison of five models show that two meta-classifier outperformed models of one factor, whether the factor was mutual reverts, bipolarity networks, or percentages of deletions among all edits. One of the meta-classifiers combined the page metrics used by Kittur et al. [104], and a list of features extracted from the collaboration networks. Jankowski-Lorek et al. [84] achieved the highest performance in predicting article controversy by combining user ratings from the Article Feedback Tool[7] and the page metrics in [104]. The combined model had an F-measure of 84.1%, a precision of 85.1%, and a ROC of 0.91.

An article's age needs to be considered in order to properly measure controversy. Frequent changes early in an article's life may not mean controversy. Instead, editors may be jointly shaping the content and navigating the tension between knowledge creation and retention [95]. Researchers have explored textual analysis of talk page discussion (e.g., presence of negative and positive words). The method was not very useful for detecting controversial articles but was useful in locating controversial parts of the articles [84].

6.4.3 Formal and informal mechanisms to resolve conflict. Wikipedia has both formal and informal mechanisms to resolve disagreements and disputes. When conflict occurs, editors typically try to resolve it via talk page discussions. Several studies have analyzed the argumentation tactics that editors use to resolve conflict on talk pages. Editors may post to compromise, agree, praise, advise, disagree, or even personally attack another editor [145]. They can also synthesize previous discussions, inform each other, ask questions, argue, or express their personal attitudes [53]. A common practice is to refer to Wikipedia policies [162], with the most highly cited policies on Consensus (CON) [193], neutral point of view (NPOV) [198], and no original research (OR) [145, 199]. Experienced editors benefited more from citing policies during talk page discussions, which reduced their likelihood of being reverted, whereas similar gestures had no effects when used by new editors ([71] Table 1).

When editors fail to resolve the disputes on their own, they can request help from other editors. The options are Third Opinion [202] to get an outside opinion, Requests for Mediation (RfM) from the Mediation Committee, which is a panel of experienced mediators, and Requests for Comments (RfCs) by posting a proposal on the article talk page and inviting comments by the broader community [83]. For example, when a disagreement is about whether to retain or remove an article, editors can invoke the Article for Deletion (AfD) process [195] and may cite reasons like impact, precedent, or relevance [132]. An editor first nominates an article for deletion and provides reasons; other editors then have seven days to state their opinions. In the end of the period, an administrator reviews the discussion and decides to keep or delete the article. Most AfD discussions focus on the four factors of notability, sources, cost of maintenance, and bias [161]. Administrators often function as mediators to help involved parties better express and respond to one another. Because Wikipedia discussions are text-based, mediators can alter the text discussion between conflicting editors (e.g., by striking through some statements), clarify ambiguity, differentiate between personal and substantive arguments, and show the editors how their exchanges could be made more constructive [19]. They can also help manage temporal discontinuities (i.e., when one party is unavailable, the other party may make misattributions), and reduce power differences among editors [19]. When conflicts get out of control, administrators can intervene by invoking

---

[7]The Article Feedback Tool (AFT) is a Wikimedia survey for article feedback, to engage readers in the assessment of article quality. Details at [190].

page protections to restrict edits to the page [79] or by banning editors. Another mechanism to prevent edit wars is the three-revert rule saying that, "an editor must not perform more than three reverts ..., in whole or in part, whether involving the same or different material ... on a single page within a 24-hour period".

How effective are these conflict resolution mechanisms? It depends on the mechanisms and type of editors involved. A key factor is the involvement of editors who are credible, experienced, and familiar with Wikipedia norms [53, 162]. These experienced and knowledgeable editors can make persuasive arguments based on precedents and use appropriate rhetoric whereas novices may formulate their arguments based on personal values and inappropriate analogies [162]. When disagreements involve experienced editors with a community orientation, such as Wikipedia administrators or editors who have contributed many edits to many articles, conflicts are less likely to escalate [10, 12]. Editor credibility can also shorten the time a discussion takes to reach consensus [94].

The mechanisms that involve polling the community, such as requests for deletions or requests for comments (RfCs) require sufficient participation to be effective. When editors fail to resolve a dispute in local discussions, they can use bots and noticeboards to publicize the dispute to a larger community and draw more participants into the discussion [83]. If all goes well, a neutral editor will summarize the discussion and make a resolution. In practice, about one third of RfCs never reached resolution mostly because of lack of participation. The lack of participation may be caused by multiple factors, such as poorly articulated statements, lack of interest from editors, long discussion threads, or excessive bickering in the discussion, which makes it hard for third parties to understand the dispute [83].

Sometimes effective conflict resolution requires not only sufficient participation, but also the right composition of editors. For example, studies of article for deletion decisions have shown that decisions made by large groups of editors with more diverse tenure were less likely to be overturned later [110]. Unfortunately, in practice, the AfD decisions are often made by "a small number of longstanding users" with only 26% of all deletion discussions including newcomers, and only 18% of discussion involving the article creators (18%) [56]. At the same time, the involvement of newcomers can be beneficial only to a certain degree. Decisions made with the participation of too many newcomers are more likely to be overturned than decisions made with fewer newcomers [110]. Part of the reason may be that newcomers struggle to understand the deletion process and/or misinterpret criteria such as notability and reliable sources [162].

6.4.4 The impact of conflict on collaboration outcomes. In contrast to the volume of research on predicting conflict and conflict resolution mechanisms, there is much less research on the impact of conflict on collaboration outcomes. The organizational behavior literature differentiates three types of conflict: task, affective, and process [87, 143]. Some researchers argue that task conflict in Wikipedia serves as "generative friction that is regulated by references to policy" [145] and is part of the coordination effort with the community to improve the quality of articles. When managed well, task conflict can improve article quality by helping to integrate diverse perspectives via "creative abrasion" [10]. In other words, task conflict can positively moderate the effects of diverse perspectives on article quality. The same study, however, shows a direct negative effect of task conflict on article quality [10]. Content analysis of talk page discussions showed that differences in editors' viewpoints could escalate into personal attacks. Although conflict is only five percent of talk page discussion, it can negatively affect article quality, especially when task conflict escalates into affective conflict or process conflict [12]. Interestingly, conflict also affects readers' perceptions of article quality. An experiment with Amazon Mechanical Workers showed that readers rated article quality lower when talk page discussions involving conflict appeared alongside with the article content [173]. When the editors resolved the conflict through compromise or collaboration, the negative effects disappeared.

Conflict negatively influences editors' willingness to contribute and the effects are greater for new and female editors than more experienced and male editors [34]. Editors who contribute to controversial articles are more likely to stop contributing to Wikipedia [20], possibly due to frustrations with long debates, personal attacks, and edit wars. Compared to male editors, female editors were more likely to cite conflict with other editors and fear of criticism as reasons for not contributing [34]. Newcomers were more likely to have their edits reverted and hence demoralized, especially when the deletions or reverts were performed by a more experienced editor or done without sufficient explanations [72]. When editors continued to contribute after being reverted, their likelihood of being reverted in the future decreased, suggesting they were learning either to be cautious or to produce better quality work [72].

How the revert is done also matters. One way of mitigating the negative effects of reverts is to provide explanations, which can be done by a human editor or by an automated quality control program. Tools like Huggle and Twinkle can expedite the speed with which vandalism gets reverted [47]), sometimes in the order of minutes [58], they also increase the chance of "good faith" newcomers getting reverted [70]. Personalized messages showing sympathy or acknowledging good faith compared to short, generic messages can mitigate the negative impact of warnings [47]. Overall, reverts have been shown as a double-edged sword. They may increase article quality by removing low-quality work and helping editors learn about community norms, but they can also drive valuable editors away [72].

Future research directions. Conflict is among the most studied aspects of Wikipedia collaboration, although most work focuses on predicting controversy and conflict resolution mechanisms. We need more research on the impact of conflict on article quality. Talk page discussion provides a rich setting to study how conflict unfolds over time and the effects of conflict on editors and article quality. Much of the coding and analysis of talk page discussion focused on classifying the rhetoric and tactics, and future research should aim to characterize the process in which conflict arises, escalates, subsides, and eventually gets resolved. How do different types of conflict manifest in talk page discussions? What tools can help detect conflict in talk page discussion and intervene before it escalates? What intelligent tools can be developed to classify types of conflict and recommend appropriate actions?

6.5 How Does Wikipedia Socialize Newcomers?

Attracting and retaining new editors is crucial to Wikipedia's long-term sustainability. As Wikipedia becomes larger and more sophisticated, newcomers struggle to understand how it works and its complex rules. Newcomer recruitment and retention became especially crucial after active editorship[8] peaked in 2007 and began declining afterward [169]. We identified 21 articles related to newcomers and their socialization. The studies have focused on the challenges faced by newcomers and the effectiveness of newcomer socialization programs, most of which were developed or implemented between 2010 and 2019.

6.5.1 Newcomer behaviors and transitions to becoming a Wikipedian. The existing literature features two schools of thought related to newcomer socialization, that is, how editors become Wikipedians or committed Wikipedia editors. According to [21], novices go through a socialization process to become Wikipedians. They start as readers visiting Wikipedia to gather information, then edit in domains of personal interest, and finally broaden participation beyond editing individual articles to contributing to the large community. In the process, they may start using tools like watchlists to track changes to articles or participate in WikiProjects. Some editors change from seeing Wikipedia as "a collective of articles with random people adding information" to seeing Wikipedia as a community; being a Wikipedian becomes part of their identity [21].

---

[8]An active editor is defined as having made five or more edits in a given month.

According to [146], Wikipedians are born, not made, with large differences between occasional editors and Wikipedians on the first day that they started participating. Most editors maintained their initial levels of contribution over time. Several other studies support this observation. For instance, survival analyses of new editors showed two clusters: occasional editors who continued to contribute to Wikipedia for less than eight hours versus customary editors who contributed for two to twenty weeks or longer [215]. There is history-dependence in editing, with early editing behaviors predicting subsequent behaviors both in terms of quantity and quality of contributions and types of activities that editors perform [4]. Editors who performed a variety of tasks in their first weeks were subsequently involved in a greater variety of tasks, such as editing, article creation, administrative activities, and vandal fighting [4].

We would like to point out that the two schools of thought do not necessarily contradict one another. Born Wikipedians, that is, editors who started with higher levels of activities, may be more likely to engage in peripheral participation by experimenting with editing articles and using tools, than editors with lower levels of activities. In the online environments, the process of peripheral participation may be compressed or expedited and hence not as visible. The key question here is not which theory wins, but how nature (natural propensity) and nurture (socialization) interact to influence editors' levels of contribution. Both studies were observational making it difficult to tease apart the influence of nature and nurture.

6.5.2   Wikipedia's resistance to newcomers and newcomers' challenges. As Wikipedia grew larger and more complicated, newcomer retention declined and became a community challenge [70, 72]. Part of the challenge is to help newcomers learn the ropes. For instance, about 29.4% of the Articles for Creation (AfC) [191] drafts created by newcomers were never submitted for review [159], possibly because newcomers did not understand the draft submission process. New editors are more likely to make edits that violate Wikipedia policies or take the article into a direction that differs from its initial one. As a result, they may have their work reverted.

Wikipedia's newcomer retention challenge reflects a universal tension in online communities between protecting an existing community and making newcomers feel welcomed and appreciated. As the number of active editors peaked in 2007, the community shifted its focus from growing the community to assuring article quality and reinforcing norms. While new editors' contribution quality remained stable over Wikipedia's history, the likelihood that good faith edits by newcomers were reverted had grown. As a result, the survival rate of newcomers had fallen substantially [70]. One study showed that the percentage of editors who made 10 or more edits and remained active a year after their joining dropped from about 40% in 2004 to about 10% in 2009 [31]. The adoption of automated tools made the matter worse. In 2011, over 80% of new editors experienced their first interaction within Wikipedia as a message from a bot, up from 40% in 2006, and the majority of these messages were notifications about what the new editors had done wrong [70].

6.5.3   Newcomer socialization programs. Newcomer socialization is the process of helping new members learn the attitudes and behaviors to adjust to their role or the environment [30]. New editors in Wikipedia face many challenges as they adapt to both the technical infrastructure and the norms of the community. They need to learn "the nuts and bolts of editing and interacting through a wiki" ([137] p. 264) and feel part of the Wikipedia community [21]. Socialization can be done in a standardized way, providing a common set of experiences for all newcomers, or in an individualized way, varying across individuals and directed by the newcomers themselves [30, 136].

Research on Wikipedia newcomer socialization has focused on two themes: analyzing existing socialization programs and developing new ones. Studies of existing programs suggest that personalized communication was more effective than standardized communication in retaining new editors. For instance, Choi et al. [30] analyzed the project talk pages and user talk pages of 50 WikiProjects and identified seven socialization tactics: invitations to join, welcome messages,

requests to work on project-related tasks, offers of assistance, positive feedback, constructive criticism, and personal comments. Although newcomers' contributions generally declined over time, receiving personalized welcome messages, assistance, and constructive criticism slowed down the decline, whereas standardized invitations to join the project led to steeper declines in newcomers' contributions.

Adopt-a-user is a socialization program created in 2006 and has had mixed success [135]. The program recruited 749 editors, who joined between 2006 and 2011 and were seeking mentors, and matched them with mentors. Communication between newcomers and their mentors predicted the amount of editing done by newcomers. Meanwhile, analyses of the communication suggests that mentors did not offer the breadth of services traditionally associated with mentorship. Mentors were largely reactive by answering questions and demonstrating their expertise, and provided limited counseling and friendship development. It was partially due to the public nature of communication and Wikipedia's policies against too much socializing (see the policy that "Wikipedia is not a blog, web hosting service, social networking service, or memorial site") [204].

In recent years, a number of socialization programs have been developed and evaluated to address the editor decline issue. Some of the programs are task-oriented, and others are socially oriented or both. Article for Creation provides a safe space for new editors to draft and submit new articles, and have the articles reviewed against minimum quality guidelines before they are published [159]. The Wikipedia Adventure is an interactive game that helps new editors accomplish seven task-oriented missions, such as setting up a profile page, communicating with other editors, making edits to articles, adding citations, etc. and earning badges for the completed missions [136]. MoodBar, introduced between 2011 and 2013, allowed new editors to share their "mood" about their first editing experience (e.g., happy, sad, confused) on a public dashboard and receive replies from a team of experienced editors [31]. The Teahouse provides socialization opportunities for new editors and creates a "safe zone" for interactive community support. New editors create a profile, introduce themselves, and have their questions answered by hosts screened for friendliness [126].

Many of these programs showed promise as well as challenges in newcomer socialization. Although editors often evaluated their experiences with the socialization programs as engaging, enjoyable, and useful, these subjectively positive experiences did not translate to positive effects on newcomer retention or contribution. For example, new editors perceived the Wikipedia Adventure game as engaging yet completing the game did not change their contribution [136]. The Article for Creation program decreased, rather than increased, new editors' productivity [159]. Use of MoodBar was associated with higher retention, yet most of the benefits occurred when a newcomer received a useful reply from existing editors [31]. Among all the new programs, Teahouse was the most effective and the only one evaluated using a controlled experiment to establish causality. Surveys of Teahouse visitor suggest that they liked the friendly atmosphere, ease of use, and the promptness and quality of the answers they received [126]. Archival analyses show that compared to non-visitors, Teahouse visitors made more edits, edited more articles, edited for longer periods, and participated more in discussion spaces, and these effects were observed for both low-activity and high-activity newcomers [126]. A controlled intent-to-treat experiment confirmed that editors who were invited to Teahouse were more like to remain active 3-4 weeks after registration and make five or more edits between 2-6 months after registration [131].

The Wiki Education Project (Wiki Ed) is another program with some success. It is an online program, started by Wikimedia Foundation, where college instructors can assign students in their classes to edit Wikipedia articles. It is a comprehensive socialization program, with practices such as tutorial and training, a clear timetable of tasks to perform, constructive guidance from Wikimedia staff, and cohort support [114]. A study comparing 16,819 new student editors in 770 Wiki Ed classes with new editors who joined Wikipedia in the conventional way shows that Wiki Ed students made more edits, performed at higher quality, and were twice as likely to continue editing Wikipedia a year after the

Wiki Ed program [114]. Besides the new programs we described above, the Wikimedia Foundation Growth Team has also experimented with other initiatives, such as simplifying account creation [180], providing feedback to new editors after they had made 10 edits [151], and providing in-context editing instructions and providing new editors suggestions about tasks they could complete [68].

6.5.4 Challenges with Wikipedia newcomer socialization. Newcomer socialization is a critical yet formidable mission for Wikipedia. What makes Wikipedia newcomer socialization so hard? The literature we reviewed suggest three tensions in newcomer socialization.

Tension between scale and personalization: An effective socialization program needs to be interactive, social, and personalized (e.g., receiving a personalized welcome message or a reply from an experienced editor) [30]. Yet personalized initiatives are difficult to scale because they require human efforts. Messages delivered by bots are highly scalable but may be ineffective or even detrimental to newcomer retention [70].

Tension between investment in human capital versus production: Both newcomers and old-timers have limited time to volunteer at Wikipedia. The time newcomers spend on training and the time that experienced editors spend on mentoring takes time away from editing articles. Laborious socialization programs like the seven missions in The Wikipedia Adventure Game reduce the time that newcomers have to edit articles [136]. Because most newcomers drop out shortly after making their first edits [126, 146], experienced editors may prefer to spend their time editing articles, which has more certain and immediate payoffs than mentoring newcomers.

Tension between visibility and psychological safety: Wikipedia sandbox illustrates the tension between community visibility and psychology safety. It provides a safe environment for newcomers to experiment with editing articles before their work is made visible to the Wikipedia community. While this safe space shields newcomers from reverts and hostile interactions, it also diminishes the impact of their work, eliminates the opportunity to receive positive feedback, and prevents them from experiencing the collaborative spirit of Wikipedia.

Future research directions. More research is needed to reconcile the two schools of thought on "Wikipedians are born, not made" versus the process of moving from peripheral participation to becoming Wikipedians [21, 113]. Wikipedia editors seem to have not strictly followed the reader-to-leader framework [152], including the progression from being a reader to becoming an editor and then a leader. Instead, editors often skip intermediate steps in their upward role transition and gaining of privileges [11]. It is possible that there is not a universal model of growth and progression for editors. New editors differ in their motivations and hence follow different trajectories to become committed Wikipedians.

Future research should continue developing socialization programs that can balance the trade-offs between scalability and personalization, human capital and production, and visibility and psychological safety. For instance, the WikiEd experiment suggests a new possibility of cohort socialization among newcomers that can be effective and scale well [114]. It can potentially address all three trade-offs. Meanwhile, there needs to be more studies of Wikipedia's anti-social culture and its impact on newcomers and the community as a whole. The existing culture explicitly discourages non-task oriented socializing among editors (e.g., encouraging only work discussion on user pages and discouraging social networking or amusement). As we mentioned in the motivations section, editors have many motivations including identity, community, and social benefits. Many new editors genuinely want to help and are capable of it. The challenge is to balance the need to nurture newcomers with the need to maintain Wikipedia as a quality encyclopedia. Some promising directions may include algorithms to identify good faith editors who are more likely to become Wikipedians and intelligent conversational agents to manage personalization and scalability.

7 WIKIPEDIA COLLABORATION EMERGENT STATES

7.1 What Roles Do Editors Assume?

In traditional organizations, roles refer to a set of behavioral expectations attached to a position in an organized set of social relationships [166]. Although there are few formal roles in Wikipedia except for administrators, research has identified various informal or emergent roles that editors can voluntarily assume. These informal roles do not necessarily constrain or enable actions; instead, they represent common behavioral patterns [9]. We reviewed 11 articles, most of which used mixed methods to classify roles, examine the dynamic evolution of roles and editors' role-taking behaviors, and the impact of editors' roles on article quality.

Research on formal roles has focused on the access that editors have to various privileges that allow them to take actions. For instance, Arazy et al. [9] examined the power relationships among editors by coding editors' privileges [203]. They identified 12 formal roles, including Benevolent Dictator (Jimmy Welsh), Directors (who provide oversight to the Wikimedia organization), Privacy Commissioners (who investigate complains about violations of privacy policy), a Security Force (who can hide page revisions and map user accounts to IP addresses), Administrators (who can delete and protect pages, block and unblock users, and edit fully protected pages), and registered users. The vast majority (99.7%) of the 4.9 million editors in 2012 were registered users who have no privileges but can edit most unprotected pages. Unregistered users are those who have not created an account and are only identified by IP address. They can edit unprotected pages but they cannot create new articles. Editors with different access privileges have different patterns of editing. For example, new editors primarily make changes to the article pages, administrators focus more on coordination work and editing talk pages, and security forces focus more on policy creation and enforcement [11].

Research on informal roles has focused on identifying the prototypical activities or common patterns of contribution that characterize each role, using either human annotators or machine learning algorithms (e.g., [6, 115, 184, 211]. Early research started with classifying editing activities (e.g., [40, 107] by analyzing the types of edits on article pages, article talk pages, user pages, user talk pages as well as structures of communication networks. For instance, Kriplean et al. [107] identified 42 activities and grouped them into six categories of editing work, social and community support, border patrol, quality control, administration, and meta-content such as building tools and templates. Daxenberger and Gurevych [40] identified a different set of task categories including creating or moving articles, adding, deleting or modifying article contents, fixing typos and grammar errors, rephrasing and restructuring existing text, adding, deleting or modifying hyperlinks, adding, deleting or modifying references, adding, deleting or modifying Wiki markup, adding or removing vandalism, etc. Using these task categories, researchers have identified a number of informal roles shown in Table 1 [6, 115, 184, 211]. We see four dominant and consistent roles from these studies.

The first role is *substantive experts* or *all-round editors* who provide substantive contents by adding information, links, references, and makeups to article pages [6, 211]. These are called all-round editors because they make many kinds of edits to article pages. Editors assuming this role also have a high rate of posting to article talk pages to explain, justify and discuss their edits [184].

The second role is *technical editors* including copy editors. Welser et al. [184] defined technical editors as those who work on incremental improvements and maintenance of Wikipedia articles. Their activities include small changes such as fixing spelling and grammar errors, rephrasing and reorganizing sentences, checking facts and links, and building templates. Other studies have split technical editors into more granular roles such as fact checkers who delete information, links, and references [211], content shapers who organize the content, and layout shapers who fix markups

Table 1. Comparison of Emergent Roles from Four Studies

Roles	Welser et al. [184]	Liu and Ram [115]	Arazy et al. [6]	Yang et al. [211]
All-round editors / substantive experts	Substantive Experts (33%)	All-Round Contributors (21.9%)	All-Round Editors (41%)	Substantive Experts
Vandal fighters / watchdogs	Vandal Fighters (6%)	Watchdogs (4.6%)	Watchdogs (13%)	Vandal Fighters
Technical / Copy editors	Technical Editors (10%)	Copy Editors (25.8%) Content justifiers (add links & refs) (28.6%) Cleaners (remove sentences, links & refs) (9.9%)	Copy Editors (12%) Content Shapers (reorganize text) (4%)	Copy Editors Fact Checkers (delete edits, links & refs) Fact Updaters (modify template &references)
Social networkers	Social Networkers (trivial percentage)			Social Networkers
Other roles		Starters (create new sentence) (9.2%)	Quick-and-Dirty Editors (add new content) (11%) Layout Shapers (makeups) (6%)	Wiki Gnomes (modify wikilinks & markup)
Data Analyzed	Edit history of 1,954 dedicated editors who joined before Jan 2004 and 5839 "new" editors who joined in January 2005	Edits on the main pages of 1, 600 articles of various quality levels in June 2009	Edits on the main pages of a stratified sample of 1,000 articles of various topics and stages in January 2012	626,761 revisions made by 38,520 editors in 172, 740 articles in December 2014

[6], and content justifiers who add links and references [115]. Technical editors typically focus most of the work on the article pages.

The third role is *watchdogs* or *vandal fighters* who specialize in fighting vandalism and blocking vandals [6, 115, 211]. Vandal fighters have moderately high levels of article page edits, followed by edits on user and Wikipedia pages because blocking vandals requires a post to the user page.

The fourth role is *social networkers* who make frequent use of communication features like article talk pages and user talk pages [184, 211]. These editors spend limited time editing articles and focus primarily on user pages, user talk pages, and Wikipedia pages where community building and support happen.

Two studies reported the distribution of editors across roles. The majority of editors were either substantive editors (22-41%), technical editors (16-64%), or vandal fighters (5-13%), with few being social networkers [6, 115, 184]. Differences in the percentages are likely due to the differences in the granular activities that define the role and different study periods. Prior studies have also uncovered additional roles such as starters who create new sentences [115], quick-and-dirty editors who focus on adding new content to articles [6], and layout shapers or Wiki Gnomes who fix markup languages [211].

Although the distribution of these roles is relatively stable, many editors change roles over time. Arazy et al. [6] referred to this phenomenon as turbulent stability, meaning the roles that an individual editor takes shift over time while the social structure of Wikipedia roles remains stable. Measuring roles as clusters of access privileges, Arazy et al. [6] found that about 34% editors changed roles, and role transitions occurred both horizontally (same-level but different privileges such as Border Patrol to Quality Assurance) and vertically (gaining or losing privileges). About 90% of role transitions were vertical, with 80% of these being upward transitions in which editors gained privileges [11]. In addition, as an article matured, the percentages of all-round contributors, content shapers, and layout shapers often declined while the percentages of quick-and-dirty editors and vandal fighters increased [6]. The work of copy editors remained stable in an article's life cycle. Researchers have also examined the roles that editors occupied across articles. The majority of editors (89.1%) assumed a single role in a single article, compared to 7.4% of assuming multiple roles in multiple articles, 3.3% of assuming a single role in multiple articles, and 0.3% of assuming multiple roles within one article (0.3%) [7].

A smaller number of studies have examined how editors occupying different roles affect article quality [211]. Contributions from substantive experts and articles dominated by all-round contributors are associated with higher quality [115, 211], because these types of editors not only insert information but also include links and references to justify the newly inserted content. For the same reason, articles dominated by content justifiers also have relatively high quality. In contrast, articles with a mixture of all-round contributors, content justifiers, and starters tend to vary in quality and articles dominated by casual contributors or starters are often of questionable quality [115]. Furthermore, the impact of roles changes as an article matures and its quality increases, with the role of substantive experts becoming less important and the role of Wiki gnomes or editors doing cleanup works becoming more important [211].

Future research directions. Research on roles has provided a good understanding of what roles have emerged, editors' role-taking behaviors, and the impact of the roles on article quality. One fruitful future direction is to understand editors' awareness and perceptions of various roles and the dynamics of editors in different roles collaborating with one another. For instance, to what degree are editors aware of each other's roles and experiences and how does such knowledge affect the way they communicate and collaborate with one another? Research on teams in traditional organizations has shown the importance of transactive memory to team effectiveness, that is, metaknowledge of who knows what and how it affects coordinate and trust among members [154]. Does similar metaknowledge exists and affects Wikipedia collaboration? Editors in different roles can employ role-specific tools, and future research can investigate how the use of these role-specific tools influence work quality (e.g., do citation tools have different impact on the work of all-around contributors vs copy editor) and improve the designs of these tools.

8 WIKIPEDIA COLLABORATION OUTPUTS

8.1 What Factors Influence Editors' Contribution?

We reviewed 38 studies that examined three streams of research related to contribution: measuring the quantity and quality of editor contribution, different types of contributions and unequal distribution of contribution across editors, and factors that affect editor contribution.

8.1.1 Measuring contribution – both quantity and quality. Two common measures of quantity are edit counts and text counts. Edit counts include the number of edits or revisions an editor makes. Text counts include the amount of text added or removed. Both measures have their limitations. Because edits vary in size, edit counts may not accurately reflect the amount of work an editor contributes. In comparison, text counts fail to capture spell checking or content

restructuring, which do not add much text but adds value. A third, relatively new measure of contribution quantity is edit sessions, which measure labor hours [57]. An edit session is a sequence of edits made by an editor without a long break (e.g., more than an hour) between edits. Long sessions indicate dedicated editing. There have been hybrid measures to account for both quantity and quality of contribution, with quality operationalized by the number of subsequent revisions over which an edit persists [1, 71, 146, 153]. The logic is that an edit that lasts a long time is of higher quality than edits that are quickly changed or reverted by subsequent editors.

There has been research to categorize types of contribution. For instance, Pfeil et al. [149] identified 13 categories of contributions such as adding information, adding link, clarifying information (e.g., rewording existing content), deleting information, deleting links, fixing links, restructuring or formatting articles, fixing grammar or spelling, marking–up text, and returning an article to an earlier state (e.g., reverting vandalism). They found the most frequent contributions were adding links and adding information. Other researchers have identified the social roles of editors who specialize in different types of contributions [184, 211]. In general, composite measures combining both quantity and quality such as total edit longevity have been shown to be robust, able to differentiate positive and negative contributions like vandalism, and reward multiple types of positive contributions [1].

8.1.2 Distribution and patterns of editor contribution.
The literature highlights several interesting patterns around editor contribution. The first is unequal contribution among editors. There is robust evidence that editors' contribution follows a power law distribution, where the majority of edits are contributed by a small group of editors, e.g., 1% of editors contribute 55% of the edits [169] or a Gini index of 0.94 showing highly uneven contributions [144]. This unequal distribution also applies to automated bots. While bots overall contribute about 15% of the content at Wikipedia, a small number of bots contribute substantially more than the long tail of other bots [62, 168].

The second pattern is the temporal change of the distribution of editors' contributions over time. The influence of elite editors (i.e., administrators or editors with more than 10,000 edits) rose after Wikipedia's launch in 2001 and then declined after 2004, as large numbers of more peripheral and casual editors joined Wikipedia. This phenomenon has been referred to as "the rise of bourgeoisie" [101]. The number of active editors and article growth both peaked in 2007 and gradually declined afterward [169]. During this period of decline, elite editors continued to invest substantive efforts, with many dedicated editors working as much as eight hours per day. This suggest that much of the decline after 2007 was a result of low recruitment and retention of the more casual editors [57].

The third pattern is the importance of an editor's contribution in the beginning of their participation on Wikipedia. Several studies have shown that editors' contributions in their first 24 to 48 hours are predictive of their long-term contribution and retention; as a result, researchers have concluded that "Wikipedians are born, not made" [146]. New editors establish their levels of contribution in their early days, with Wikipedians (i.e., people who become heavy contributors) showing significantly higher quantity and quality of contribution in their early days than non-Wikipedians (i.e., those who become casual editors [4, 146, 215]. The casual editors typically only contributed a day or less before disappearing whereas Wikipedians continued contributing for two to twenty weeks or longer [215]. This pattern is also true for the types of contribution. Editors who performed a variety of tasks in their first weeks were involved in a greater variety of tasks, such as editing, article creation, administrative activities, and vandal fighting, over time [4].

8.1.3 Factors that affect editor contribution.
A third research stream focuses on factors that affect editor contribution, such as editor characteristics, feedback, identification with Wikipedia, connections with other editors, and external events. Editor motivation is also a key factor that affects contribution and we addressed it in the section on motivation. Among the various factors, two are especially worth noting.

One factor is editors' experiences with the Wikipedia community, such as the feedback they receive, their affiliation with WikiProjects, and their social connections with other editors [220]. Although receiving feedback should increase contribution in theory, the actual effects are nuanced and contingent upon the type of feedback and the editors' experience. Editors who received negative feedback (e.g., warnings and reprimands) or directive feedback (e.g., instructions, assigned tasks, or goals) contributed more to the articles on which they received the feedback but this boost in contribution did no generalize to other articles. In contrast, editors who received positive feedback (e.g., acknowledgements and rewards) or social feedback (e.g., social support or personal connections) contributed more to Wikipedia in general [220]. Interestingly, these differential effects of feedback occurred only to relatively new editors but not to more experienced editors. WikiProject affiliation also influences editors' contribution. Joining a WikiProject slightly increased editors' edits to Wikipedia as a whole, and significantly increased their contributions to the articles within the scope of the project [103]. The impact also depends on the nature of editors' pre-existing connections with the project. Editors with either identity-based attachment to a WikiProject (i.e., having interests aligned with the topics of the WikiProject) or bonds-based attachment (having interacted with other project members) had higher levels of contribution and lower likelihood of withdrawal; bonds-based attachments had stronger effects than the identity-based ones [213].

Another factor that may influence editor contribution to certain articles or topics is the occurrence of current, external events, such as breaking news, natural disasters, sports game results, the death of famous people, and ongoing social movements, all of which can drive contributions to the articles covering these topics (e.g., [97, 112]). Intense co-editing activities often happened right after the occurrence of an external event and then decayed over time. This pattern has been shown across multiple events, such as the "tohoku earthquake and tsunami" [97], the "Egyptian revolution" [48], "2012 deaths" [99], and "Black Lives Matter" [176]. The spike in editing occurred on both article main pages and talk pages [48]. Another interesting pattern is that editors working on these current-event articles collaborated across articles and not just within a single article [97–99]. For example, following the Trayvon Martin shooting, which galvanized the Black Lives Matter (BLM) movement, 50% of editors who contributed to the top-10 articles in the category also edited non-top 10 articles. They also visited old BLM articles while working on newly created ones [176]. The impact of external current events is not always positive. When football teams lose, their fans reduce contribution to football-related pages on Wikipedia [112], and the reduction is greater when the loss is by a larger margin or to a rival team.

Future research directions. We see two future directions: a macro one and a micro one. At the macro level, we need more longitudinal research of the evolution of Wikipedia as a whole as well as editor behaviors and various collaboration processes. Similar to organizations and other social entities, online collaboration projects like Wikipedia have their own life cycles and go through different stages of growth and decline [169]. Editors behaviors and interaction dynamics may differ across the different stages. We need more research to identify and characterize the various life stages and more studies like Arazy et al. [6], which examines the dynamics of editors assuming different roles over time. At the micro level, we need more research on how editors choose their contribution types and tasks, i.e., how editors decide which articles to edit and the types of work they perform. Prior research has shown great variety in editors' contribution patterns, with most editors contributing to a single or a few articles, a subset of editors contributing to many articles, and the most central editors making many contributions to many articles [97]. How do editors decide which articles to edit and which tasks to perform? Is the article choice a random process or guided by heuristics or tools? These questions are important to understand because they provide insights about why certain tasks or topics may struggle

to attract editors. The insights can also provide implications for the design of tools to match editors with tasks and encourage future contribution.

8.2 What Predicts Article Quality?

Quality is the most frequently studied area of Wikipedia research. We classified 36 papers in this category, most of which centered around measuring the quality of edits or articles and explaining what leads to high quality.

8.2.1 Measuring and predicting article quality. Quality is a multi-dimensional concept in the context of Wikipedia articles. High article quality requires accuracy of and up-to-date information presented with a neutral point of view, completeness of coverage, and comprehensibility of writing [124]. Article quality has been assessed using both manual and automated methods. Manual methods include expert and crowdsourced ratings (e.g., [85]), and typically can achieve better accuracy than automated methods [171]. One example of expert ratings is WikiProject Editorial Team Assessment [194], in which editors manually review and assign articles to one of seven quality classes (Stub, Start, C-class, B-class, Good Article, A-class, and Featured). As Wikipedia grew larger, manual methods were difficult to scale and barely kept up with the growth in the number of articles and changes in them.

Several studies proposed automated/machine-learning methods of determining article quality, generally using the Editorial Team Assessment's manual ratings as the ground truth. The automated methods used several classes of features to predict article quality, including article attributes like length, process features like reverts or persistence of the text in an article, editor attributes, interactions among the editors, and editing patterns over the article's lifecycle. Article length is included in most automated measurements and is highly associated with human judgments of article quality [81]. Other article attributes include the number of headings as a signal of the article structure and the number of references, images, and Wikilinks to assess an article's completeness [183]. Process features such as the number of reverts, number of revisions per editor, and percentages of anonymous users, improved quality predictions and were more accurate in predicting the quality of controversial articles than models with only article attributes [42]. Persistent edits (i.e., ones not changed or deleted by subsequent editors) are an indicator of higher quality work, and the number of persistent edits made to an article are a better predictor of article quality than the number of transient edits [209].

Several editor attributes predict both article quality and the quality of individual edits, such as the number of editors who have contributed to an article and the quantity and quality of their work. High-quality articles had twice as many editors as low-quality articles [43]. Editors' prior levels of contribution and experience were also important in predicting article quality. Both the proportion of articles to which the editors had contributed that were of high quality and the number of edits they made predicted high-quality articles [18]. The quality of editors' prior contribution also matters. High-quality editors tended to produce more long-lived contributions [37]. Using the persistence of an editor's edits to measure editor "authority," Hu et al. [81] show that the authority of the editors who introduced contents predicted article quality. In addition, the quality of reviewers, i.e., editors who read the original content without changing it, also predicted article quality.

Prior collaboration among editors is another useful predictor of article quality, and co-editing networks show the degree to which editors have previously co-edited articles [43]. The existence of a link between two editors indicates they had both edited at least one article and the edge weight indicates the number of articles they had co-edited. The co-editing networks of high-quality articles were both denser and had higher weights than the networks of low-quality articles, implying that the articles were more likely to be of high quality if more of the contributing editors had collaborated in the past and if the collaborations were more frequent [43].

8.2.2 *Explaining article quality.* Many factors can influence article quality, including the number of editors, their levels of contribution, their expertise on the article's topic, and coordination among editors. Some of these factors overlap with those we mentioned earlier in measuring and predicting article quality. We mentioned some of the factors in earlier sections on coordination and roles. In this section, we try to consolidate and recap all into one summary.

The first set of factors includes the *number of edits and editors*. Featured Articles have a larger number of edits made by more unique editors than non-Featured Articles, after controlling for article's age and popularity [207]. Factors such as article visibility can influence the number of edits and editors, and visible articles tend to attract more editors and edits. Hence, a small number of articles on topics of high interest to the Wikipedia community can attract a disproportionately number of edits [207]. Some studies have shown a curvilinear relationship between the number of editors and article quality, with additional contributors increase quality up to a point and then more contributors detract from the quality [96].

The second factor is the breadth of the *editors' experience*, which can be measured by the number of articles to which an editor has contributed. Editors' breadth of expertise from editing different articles and topics can help improve article quality [105]. Interestingly, Kane and Ransbotham [96] studied the affiliation networks among article of the Medicine WikiProject and found that articles that shared editors with high-quality articles tend to have higher quality, partially because the diverse knowledge and experiences of the editors improved quality.

The third factor is the *roles* that editors play, which represent the type of work they contribute to an article. In particular, articles that are dominated by all-round contributors and substantive editors tend to be of higher quality than articles dominated by casual contributors and protectors who revert vandalism [115]. Articles dominated by copy-editors are also of disproportionately high quality [115].

The fourth factor is *coordination* among editors. Coordination may occur within articles (e.g., via discussion on talk pages) or across articles (via WikiProject discussion). High-quality articles have more intense coordination via discussion pages than low-quality articles [207]. Adding more editors to an article is only beneficial to the extent that they can effectively coordinate their work, e.g., by having a few editors structure an article before others pitch in to help expand it [102]. This form of coordination, referred to as implicit coordination or editing inequality, is positively associated with article quality [8, 102]. In comparison, discussion via talk pages, referred to as explicit coordination, is more effective when fewer editors are contributing [102].

There have been multiple quality-improvement interventions to mobilize editors to improve Wikipedia articles, such as the "Collaborations of the Week," which organizes WikiProject members to improve a small set of designated articles, and WikiCup contests, in which editors score points for achieving specific tasks. Surprisingly, although these interventions improved article quality, a greater number of contributors was associated with a smaller increase in quality [105, 182]. One possible reason for the negative association may be the difficulty of coordinating large groups of contributors. Simply drawing attention to quality improvement projects do not guarantee quality improvement without clear incentives and task structure. Interventions where editors work individually or in small groups were more effective than interventions where editors work in large groups [182].

The fifth factor is the prior *collaboration network* among editors working on an article. A link between two editors may indicate either they had previously contributed to the same article or that they had exchanged messages via user pages. Measuring prior collaboration as interactions on user pages, Nemoto et al. [138] showed that more centralized and cohesive collaboration networks among editors increased both the likelihood and the speed of article promotion to Good Article or Featured Article status. Greater centralization of the collaboration networks implies the existence

of an editor connected with many other editors, and greater cohesiveness implies the existence of small clusters of interconnected editors. Article quality is associated with more centralized and cohesive collaboration networks.

The sixth factor is the *timing* of edits. High-quality articles often experienced a period of highly intensive editing before they acquired either Good Article or Featured Article status [209]. Featured Articles on average contained high–quality content approximately 86% of the time over their life cycles and this value increased to 99% during the last 50 revisions of the articles [86]. So not only did high-quality articles receive more edits from editors over their development, they also tended to quickly converge in periods of intense editing to a high-quality final version.

8.2.3 Vandalism – a significant threat to article quality. We found two articles related to vandalism and include them here because vandalism is the antithesis of quality work. The first paper is a literature review of 67 research papers on vandalism in Wikipedia that were published from 2007 to 2015 [174]. Almost 60% of the research reviewed aimed to develop automated methods to detect vandalism (40 articles), followed by research on "the less important quality control of Wikipedia content (12% of articles), content analysis (6% articles), and prevention (3% articles). Unfortunately, for the purposes of the current paper, Tramullas et al. [174] did not reach any substantive conclusions about the relationship between vandalism and article quality. The second paper examined different types of vandalism and associated damage [153]. They identified seven types of damage, such as misinformation, mass deletion, partial deletion, offensive language, spam, nonsense, and others. Damage was assessed as the number of views on the article before the vandalism was repaired. Approximately 42% of damage was fixed quickly, within one article view, and other incidents took much longer, e.g., up to 35 hours or even 3 months. The impact of vandalism also depended on how often it occurred and how hard it was to detect automatically. For example, mass deletion is easy to detect and can therefore be easily repaired. In comparison, insertion of offensive content and misinformation are considered highly damaging because they are common, hard to detect, and detrimental to article quality.

Future research directions. There are two promising future directions for article quality. The first direction is the development of automated tools to both predict article quality and to identify areas to improve articles, e.g., missing links or problematic writing. This requires making the prediction algorithms transparent so that they can reveal the factors that are associated with low article quality and hence that need improvement. The second direction is to understand the motivations behind the people who vandalize article. For example, why do vandals engage in their damaging behaviors and is it possible to prevent these actions rather than repair them after the fact? Most research has focused on reacting to vandalism by reverting the changes. By understanding the motivations of vandals, researchers could explore a more proactive approach to address the issue.

9 GENERAL DISCUSSION

Table 2 summarizes the key insights from our literature review and promising directions for future research. In the rest of this section, we discuss some of the insights in light of traditional theories of off-line social behavior and organizations and identify ways to expand these traditional theories to account for the newer contexts of online collaboration. We focus on the topics of motivation, coordination, leadership, conflict, and newcomer socialization, all of which have been extensively studied in the context of traditional organizations. We also discuss the implications of the review for the design of collaborative and social computing systems and Wikipedia research across disciplines.

In terms of motivation, the three types of motivations – intrinsic, extrinsic, and social – correspond to motivations identified in the job motivation literature [65, 66, 157]. Some of the findings seem to be consistent with insights from the motivation literature. For example, the discrepancy between what editors say that motivates them (fun and ideology)

Table 2. Summary of Key Insights and Future Directions

Key Questions (# of articles reviewed)	Key Insights / Takeaways	Future Directions
Who are the **editors**? (27)	• Young, male, computer savvy, time online. • Few (10-18%) female editors. • Gender gap caused by a hostile culture with conflict and reverts. • Under-representation of women leads to less female-oriented content and bias in how men and women are portrayed in articles.	• Study the impact of female editors on collaboration processes and article quality. • Develop and evaluate interventions to recruit and support female editors.
What **motivates** the editors? (12)	• A variety of motivations: intrinsic (fun and learning), extrinsic (reputation), and social (altruism, ideology, sense of community). • Contribution more often driven by self-interest, recognition, and social image, not altruism or ideology. • Motivations change as editors gain experiences and focus more on social and community benefits.	• Need more up-to-date research on editor motivations now that Wikipedia is established. • Need theory to explain how motivations change over the life cycle of editors and Wikipedia. • Need research on how motivations affect different types of contributions beyond article editing work.
How do **bots** affect the work of editors? (16)	• Different types of bots (generator, fixer, connector, archiver, protector, advisor, etc.). • Advisor and task suggestion bots positively affect editor's contribution. • Vandal fighting bots improve article quality but negatively effect contribution. Negative effect is less negative with a longer, more friendly and informative message.	• Need to study human awareness, perception, and acceptance of bots. • Need to study complex interactions between human editors and bots. • Need to study algorithm bias and value-sensitive designs.
How do editors **coordinate** their work? (29)	• Both explicit coordination and implicit coordination improve article quality. Explicit coordination is less effective for articles with many editors than implicit coordination. • Talk page discussions dominated by requests for editing, requests for information, and references to guidelines and policies. • Editors who join WikiProjects shift to project-related work, maintenance work, and interactions with project members.	• Need to develop tools to manage task interdependencies (e.g., reminders to post on talk pages before making changes). • Need to study social dynamics on talk pages (e.g., how talk page discussions evolve and how new editors make sense of talk page discussions).
What are the **governing** rules and policies? (13)	• Rule-based governance occurs though policies, guidelines, essays. • These mechanisms are complex, with ambiguities around interpretations of policies and interplay of multiple policies. • Algorithmic governance with bots is scalable and key for vandal fighting but harms participation of new editors and biased against anonymous and new editors.	• Need to understand the temporal evolution of policies. • Need to understand the processes through which newcomers learn about policies and develop tools to help new editors navigate the policies. • Need research on essays (e.g., their creation and influence).

Key Questions (# of articles reviewed)	Key Insights / Takeaways	Future Directions
Who are the **leaders**? How do they lead? (13)	• Formal roles (e.g., administrators) and shared leadership by ordinary editors. • Admins success depends on # of edits, diverse edits, and edit summary. Harder to become an admin after 2006 with greater emphasis on policy and WikiProject edits. • Positive feedback, directing, and social exchanges from peer editors increase contribution; negative feedback reduces it.	• Need more research on what happens after editors attain admin status. • Need richer analysis of communication strategies used by admins and ordinary editors to influence other editors.
How do editors manage and resolve **conflict**? (34)	• Conflict arises from both article topics and the behaviors of editors. • Conflict negatively affects editors, especially newcomers and women. • Controversy prediction requires combining multiple factors, esp. data from talk pages and editors' collaboration networks. • Effective conflict resolution requires community participation, esp. editors who are familiar with Wikipedia norms.	• Need research on the effects of conflict on individual editor experience and article quality. • Need research on the process of conflict escalation and resolution, e.g., content analysis of talk page discussion. • Need to develop tools to intervene when conflict arises.
How are **newcomers socialized** to become Wikipedians? (21)	• Newcomers are more likely to be reverted. Being reverted is demotivating, and a personalized message can mitigate the effect. • Personalized tactics are more effective than standardized tactics, esp. a personalized message from experienced editors. • Most existing socialization programs failed to increase retention and contribution, except for Teahouse and the WikiEd Project.	• Need to further reconcile Wikipedians born or made debate. • Need to develop more effective socialization programs, to balance trade-offs between scalability and personalization, newcomer socialization and article production, visibility and psychological safety. • Develop AI tools to identity and train good faith newcomers.
What **roles** do editors assume and how do they operate? (11)	• Four common roles: all-round editors / substantive experts, technical / copy editors, vandal fighters, and social networkers. • Articles dominated by all-round contributors and technical editors achieve higher quality; importance of roles changes with articles' lifecycle. • Roles at the system level are stable although many individual editors change their roles over time.	• Need research on how editors in different roles perceive and collaborate with one another. • Need research on how role-based behaviors influence article quality. • Need to develop role-specific tools and study their impact on collaboration.
What factors influence editors' **contribution**? (38)	• Measures of both quantity (edit or word counts) and quality of contribution (how long edits persisted through revisions). • Contributions per editor showed a power law distribution and a decline in edits by elite editors', i.e., "the rise of bourgeoisie."	• Need more research on the longitudinal evolution of Wikipedia and its life cycle. • Need more research on how editors choose articles and tasks, e.g., why some editors work on few articles / tasks and others work on many.

Key Questions (# of articles reviewed)	Key Insights / Takeaways	Future Directions
	• Quantity of contributions influenced by positive and negative feedback, WikiProject affiliation, and external events.	
What is edit and article **quality**? What inputs and processes affect quality? (36)	• Factors in machine learning models to predict quality: article length, # of references, # of headings, controversy, # of editors, editors' past contribution, and social networks among editors. • Variables that explain quality: # of edits, # of editors, editor experiences and roles, coordination, editors' relationships. • Types of vandalism damage: mass deletion, partial deletion, misinformation, offensive, spam, nonsense. Most vandalism is fixed quickly, 42% within one article view.	• Need to develop more transparent article assessment tools to identify specific parts of an article that need improvement. • Need to understand motivations for vandalism in order to develop interventions to proactively identify and prevent vandalism behaviors.

and what actually motivates them (self-benefit and recognition) may be due to the "crowding out" effect of extrinsic motivation on intrinsic motivation [44]. Meanwhile, the voluntary nature of Wikipedia participation makes it a good setting to study individual differences in motivations, how individual motivations evolve over time, and the interplay among various motivations when none dominates the others.

In terms of coordination, a key difference between Wikipedia and traditional organizations is the lack of managerial oversight to decide who does what, when, where, and how. Instead, coordination in Wikipedia relies primarily on editors' ability to see the current state of an article and their individual decisions about what voluntary actions they should take to improve the product. Our review highlights a few coordination mechanisms to facilitate communication among editors and to match editors to various articles and tasks (e.g., WikiProject and Collaborations of the Week). To improve coordination among editors, we can borrow insights from the traditional coordination literature. For example, common ground and shared mental models can lead to positive team processes and outcomes (e.g., [32, 119]). Future research could explore the development of tools to facilitate common ground and shared knowledge among editors, e.g., providing new editors contextual information related to the section they are editing. Automated Tools can also help monitor the number of editors who work on an article at any time and provide recommendations or warnings about potential coordination conflict, costs or pitfalls.

In terms of leadership, a key difference between Wikipedia and traditional organizations is the importance of informal shared leadership among editors. Although the traditional literature on leadership in off-line organizations has acknowledged the role of shared leadership [221], it primarily focuses on traits of good leaders, different types of leadership styles such as transactional and transformational, leader-member exchanges, etc. In contrast, the Wikipedia literature focuses more on leadership as a behavior, instead of status or power. The traditional leadership literature can benefit from research on Wikipedia, especially in understanding the member behaviors associated with shared leadership and their effectiveness [106, 221].

The traditional literature on conflict has focused on various types of conflict (e.g., task conflict or relationship conflict), the impact of conflict on team performance, and conflict resolution theories such as communication strategies and mediation to resolve conflict. In contrast, the Wikipedia literature has focused on measuring and predicting conflict

as well as mechanisms for conflict resolution. One common pattern across the two contexts is the negative impact of conflict – both task and relationship conflict [41]. Research in traditional teams also suggests that the negative association between task conflict and team performance is weaker when task conflict and relationship conflict are weakly, instead of strongly, correlated [41]. Hence, task conflict that does not escalate to relationship conflict causes less harm. A great benefit of studying conflict in Wikipedia is that compared to traditional organizations, Wikipedia maintains a mostly complete record of the communication and work artifacts associated with the conflict. This allows for a deeper and more dynamic study of the sources, escalation and consequences of conflict. On the other hand, organizations can learn from Wikipedia in using automated tools to identify potential conflict, by analyzing electronic communication such as emails, instant messaging, and video conferencing transcripts. Automated tools can use the information to provide coaching and guidance to employees.

The traditional literature on newcomer socialization has examined socialization as both formal and informal learning and as a long-term process with various stages, contents, and effects. In traditional organizations, there are different socialization tactics such as institutionalized (collective and sequential content) versus individualized (individual and random content) [88]. Successful socialization occurs if the newcomer becomes committed and satisfied with organizational membership, conforms to its culture, and performs well. Some of the tactics may shed light on Wikipedia socialization, such as having newcomers go through socialization collectively, instead of individually. This may explain the success of the Wiki Education Project in which students go through socialization in cohorts [114]. Meanwhile, our review suggests that newcomers often do not have time to go through a lengthy socialization process. The three trade-offs we identified between scalability and personalization, human capital and production, visibility and psychological safety hopefully can inform development of future socialization programs.

Most of the research we reviewed focused on Wikipedia as a self-contained platform and did not consider the ecosystem in which Wikipedia operates, especially the rise of popular social media platforms such as Facebook, Twitter, YouTube, Instagram, and many others. Most of these platforms started around 2004-2006, one or two years preceding the decline in the number of Wikipedia editors. Besides the popularity of bots, the rise of many other social media platforms may be another important reason behind Wikipedia's decline. The organizational ecology literature [74, 181] and the research on platform competition [156] both provide valuable insights to better understand external forces that might have contributed to the decline. As an early mover among the social media platforms, Wikipedia enjoyed an early mover advantage for several years and attracted a critical mass of contributors. The emergence of later and more popular social media platforms competed with Wikipedia for a limited contributor base [181], which exacerbated the challenge of recruiting and retaining new editors. However, these thoughts about the role of the social media ecosystem are only speculation and rigorous empirical studies are needed to test and confirm our proposition.

Our review highlights several key lessons for building successful online collaboration projects. It is important to appeal to multiple motivations, e.g., opportunities for learning, helping, and socializing, and to attract volunteers with diverse backgrounds. It is important to have effective governance, leadership, and mechanisms to coordinate collective efforts, socialize newcomers, and resolve conflict. In an era of rampant disinformation on other social media, Wikipedia is able to remain a trusted information source because of shared values among many editors (e.g., neutral point of view), policies requiring information come from trusted sources, a social and technical infrastructure that allows individual editors to identify and repair errors, and policies and governance structures that help editors handle inevitable conflict. It is also important to deploy tools such as bots to augment human efforts, although such tools may have unintended consequences and need to be designed around human values.

In addition, we have two high-level observations about the challenges of designing collaborative and social computing systems in general. The first challenge is the tension and interplay between system designs and emergent behaviors, which jointly shape the processes and outcomes of Wikipedia collaboration. In other words, collective human behaviors are driven not only by system designs (e.g., talk pages and administrator roles) but also by the emergence of autonomous human behaviors (e.g., discussions on the talk pages and shared leadership). The goal of designing such systems should not be to control human behaviors, but rather to enable and facilitate desirable human behaviors. The second challenge is the trade-offs involved in making design decisions, such as balancing the goals of protecting Wikipedia from vandalism and misinformation with the goal of socializing and retaining new editors.

Finally, although the primary focus of our review is on editors' behavior and collaboration processes, there are many other important issues related to Wikipedia, such as understanding reader behaviors and consumption dynamics by analyzing page views and click stream data, exploring the roles of links and images in collaboration processes, and using Wikipedia articles as training data for large language models. While these issues are beyond the scope of our work, they deserve their own reviews in future research.

ACKNOWLEDGMENTS

REFERENCES

[1] B Thomas Adler, Luca De Alfaro, Ian Pye, and Vishwanath Raman. 2008. Measuring author contributions to the Wikipedia. In *Proceedings of the 4th International Symposium on Wikis*. 1–10.

[2] Yann Algan, Yochai Benkler, Mayo Fuster Morell, and Jérôme Hergueux. 2013. Cooperation in a Peer Production Economy Experimental Evidence from Wikipedia. *Available at SSRN 2843518* (2013).

[3] Judd Antin. 2011. My kind of people? Perceptions about Wikipedia contributors and their motivations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 3411–3420.

[4] Judd Antin, Coye Cheshire, and Oded Nov. 2012. Technology-mediated contributions: editing behaviors among new wikipedians. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 373–382.

[5] Judd Antin, Raymond Yee, Coye Cheshire, and Oded Nov. 2011. Gender differences in Wikipedia editing. In *Proceedings of the 7th international symposium on wikis and open collaboration*. 11–14.

[6] Ofer Arazy, Johannes Daxenberger, Hila Lifshitz-Assaf, Oded Nov, and Iryna Gurevych. 2016. Turbulent stability of emergent roles: The dualistic nature of self-organizing knowledge coproduction. *Information Systems Research* 27, 4 (2016), 792–812.

[7] Ofer Arazy, Hila Liifshitz-Assaf, Oded Nov, Johannes Daxenberger, Martina Balestra, and Coye Cheshire. 2017. On the" how" and" why" of emergent role behaviors in Wikipedia. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 2039–2051.

[8] Ofer Arazy and Oded Nov. 2010. Determinants of wikipedia quality: the roles of global and local contribution inequality. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 233–236.

[9] Ofer Arazy, Oded Nov, and Felipe Ortega. 2014. The [Wikipedia] World is not flat: on the Organizational Structure of Online production Communities.. In *ECIS*.

[10] Ofer Arazy, Oded Nov, Raymond Patterson, and Lisa Yeo. 2011. Information quality in Wikipedia: The effects of group composition and task conflict. *Journal of management information systems* 27, 4 (2011), 71–98.

[11] Ofer Arazy, Felipe Ortega, Oded Nov, Lisa Yeo, and Adam Balila. 2015. Functional roles and career paths in Wikipedia. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 1092–1105.

[12] Ofer Arazy, Lisa Yeo, and Oded Nov. 2013. Stay on the Wikipedia task: When task-related disagreements slip into personal and procedural conflicts. *Journal of the American Society for Information Science and Technology* 64, 8 (2013), 1634–1648.

[13] Phoebe Ayers. 2006. Researching wikipedia-current approaches and new directions. *Proceedings of the American Society for Information Science and Technology* 43, 1 (2006), 1–14.

[14] Judit Bar-Ilan. 2018. Tale of three databases: The implication of coverage demonstrated for a sample query. *Frontiers in Research Metrics and Analytics* 3 (2018), 6.

[15] Judit Bar-Ilan and Noa Aharony. 2014. Twelve years of Wikipedia research. In *Proceedings of the 2014 ACM conference on Web science*. 243–244.

[16] Hoda Baytiyeh and Jay Pfaffman. 2009. Why be a Wikipedian. In *Proceedings of the 9th international conference on Computer supported collaborative learning-Volume 1*. 434–443.

[17] Daryl J Bem. 1995. Writing a review article for Psychological Bulletin. *Psychological Bulletin* 118, 2 (1995), 172.

[18] Grace Gimon Betancourt, Armando Segnine, Carlos Trabuco, Amira Rezgui, and Nicolas Jullien. 2016. Mining team characteristics to predict Wikipedia article quality. In *Proceedings of the 12th International Symposium on Open Collaboration*. 1–9.

[19] Matt Billings and Leon A Watts. 2010. Understanding dispute resolution online: Using text to reflect personal and substantive issues in conflict. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1447–1456.

[20] Ulrik Brandes, Patrick Kenis, Jürgen Lerner, and Denise Van Raaij. 2009. Network analysis of collaboration structure in Wikipedia. In *Proceedings of the 18th international conference on World wide web*. 731–740.

[21] Susan L Bryant, Andrea Forte, and Amy Bruckman. 2005. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. 1–10.

[22] Alan Bryman. 2013. *Leadership and organizations (RLE: Organizations)*. Routledge.

[23] Luciana S Buriol, Carlos Castillo, Debora Donato, Stefano Leonardi, and Stefano Millozzi. 2006. Temporal analysis of the wikigraph. In *2006 IEEE/WIC/ACM International Conference on Web Intelligence (WI 2006 Main Conference Proceedings)(WI'06)*. IEEE, 45–51.

[24] Moira Burke and Robert Kraut. 2008. Mopping up: modeling wikipedia promotion decisions. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 27–36.

[25] Moira Burke and Robert Kraut. 2008. Taking up the mop: identifying future wikipedia administrators. In *CHI'08 extended abstracts on Human factors in computing systems*. 3441–3446.

[26] Brian Butler, Elisabeth Joyce, and Jacqueline Pike. 2008. Don't look now, but we've created a bureaucracy: the nature and roles of policies and rules in wikipedia. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 1101–1110.

[27] Stevie Chancellor and Munmun De Choudhury. 2020. Methods in predictive techniques for mental health status on social media: a critical review. *NPJ digital medicine* 3, 1 (2020), 43.

[28] Kathy Charmaz. 2014. *Constructing grounded theory*. sage.

[29] Yan Chen, Rosta Farzan, Robert Kraut, Iman YeckehZaare, and Ark Fangzhou Zhang. 2020. Motivating Experts to Contribute to Digital Public Goods: A Personalized Field Experiment on Wikipedia. *Management Science (In press)* (2020).

[30] Boreum Choi, Kira Alexander, Robert E Kraut, and John M Levine. 2010. Socialization tactics in wikipedia and their effects. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 107–116.

[31] Giovanni Luca Ciampaglia and Dario Taraborelli. 2015. MoodBar: Increasing new user retention in Wikipedia through lightweight socialization. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 734–742.

[32] Herbert H Clark. 1996. *Using language*. Cambridge university press.

[33] Noam Cohen. 2011. Define gender gap? Look up Wikipedia's contributor list. *The New York Times* 30, 01 (2011).

[34] Benjamin Collier and Julia Bear. 2012. Conflict, criticism, or confidence: An empirical examination of the gender gap in Wikipedia contributions. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 383–392.

[35] Dan Cosley, Dan Frankowski, Loren Terveen, and John Riedl. 2007. SuggestBot: using intelligent task routing to help people find work in wikipedia. In *Proceedings of the 12th international conference on Intelligent user interfaces*. 32–41.

[36] Kevin Crowston, Kangning Wei, James Howison, and Andrea Wiggins. 2008. Free/Libre open-source software development: What we know and what we do not know. *ACM Computing Surveys (CSUR)* 44, 2 (2008), 1–35.

[37] Alberto Cusinato, Vincenzo Della Mea, Francesco Di Salvatore, and Stefano Mizzaro. 2009. QuWi: quality control in Wikipedia. In *Proceedings of the 3rd workshop on Information credibility on the web*. 27–34.

[38] Linus Dahlander and Siobhan O'Mahony. 2011. Progressing to the center: Coordinating project work. *Organization science* 22, 4 (2011), 961–979.

[39] Sanmay Das, Allen Lavoie, and Malik Magdon-Ismail. 2016. Manipulation among the arbiters of collective intelligence: How Wikipedia administrators mold public opinion. *ACM Transactions on the Web (TWEB)* 10, 4 (2016), 1–25.

[40] Johannes Daxenberger and Iryna Gurevych. 2013. Automatically classifying edit categories in Wikipedia revisions. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. 578–589.

[41] Carsten KW De Dreu and Laurie R Weingart. 2003. Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis. *Journal of applied Psychology* 88, 4 (2003), 741.

[42] Gabriel De la Calzada and Alex Dekhtyar. 2010. On measuring the quality of Wikipedia articles. In *Proceedings of the 4th workshop on Information credibility*. 11–18.

[43] Baptiste de La Robertie, Yoann Pitarch, and Olivier Teste. 2015. Measuring article quality in Wikipedia using the collaboration network. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. 464–471.

[44] Edward L Deci, Richard Koestner, and Richard M Ryan. 1999. A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological bulletin* 125, 6 (1999), 627.

[45] Katherine Ehmann, Andrew Large, and Jamshid Beheshti. 2008. Collaboration in context: Comparing article evolution among subject disciplines in Wikipedia. *First Monday* (2008).

[46] Rosta Farzan, Saiph Savage, and Claudia Flores Saviaga. 2016. Bring on board new enthusiasts! A case study of impact of Wikipedia art+ feminism edit-a-thon events on newcomers. In *Social Informatics: 8th International Conference, SocInfo 2016, Bellevue, WA, USA, November 11-14, 2016, Proceedings, Part I 8*. Springer, 24–40.

[47] Ryan Faulkner, Steven Walling, and Maryana Pinchuk. 2012. Etiquette in wikipedia: Weening new editors into productive ones. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. 1–4.

[48] Michela Ferron and Paolo Massa. 2011. Collective memory building in Wikipedia: the case of North African uprisings. In *Proceedings of the 7th international symposium on wikis and open collaboration*. 114–123.

[49] Claudia Flores-Saviaga, Saiph Savage, and Dario Taraborelli. 2016. Leadwise: using online bots to recruit and guide expert volunteers. In *Proceedings of the 19th ACM Conference on Computer Supported Cooperative Work and Social Computing Companion*. 257–260.

[50] Heather Ford and R Stuart Geiger. 2012. " Writing up rather than writing down" becoming Wikipedia literate. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. 1–4.

[51] Andrea Forte, Niki Kittur, Vanessa Larco, Haiyi Zhu, Amy Bruckman, and Robert E Kraut. 2012. Coordination and beyond: social functions of groups in open content production. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*. 417–426.

[52] Andrea Forte, Vanesa Larco, and Amy Bruckman. 2009. Decentralization in Wikipedia governance. *Journal of Management Information Systems* 26, 1 (2009), 49–72.

[53] Dominique Fréard, Alexandre Denis, Françoise Détienne, Michael Baker, Matthieu Quignard, and Flore Barcellini. 2010. The role of argumentation in online epistemic communities: the anatomy of a conflict in Wikipedia. In *Proceedings of the 28th Annual European conference on cognitive ergonomics*. 91–98.

[54] Jana Gallus. 2017. Fostering public good contributions with symbolic awards: A large-scale natural field experiment at Wikipedia. *Management Science* 63, 12 (2017), 3999–4015.

[55] R Stuart Geiger. 2009. The social roles of bots and assisted editing programs. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*. 1–2.

[56] R Stuart Geiger and Heather Ford. 2011. Participation in Wikipedia's article deletion processes. In *Proceedings of the 7th international symposium on Wikis and open collaboration*. 201–202.

[57] R Stuart Geiger and Aaron Halfaker. 2013. Using edit sessions to measure participation in Wikipedia. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 861–870.

[58] R Stuart Geiger and Aaron Halfaker. 2013. When the levee breaks: without bots, what happens to Wikipedia's quality control processes?. In *Proceedings of the 9th International Symposium on Open Collaboration*. 1–6.

[59] R Stuart Geiger and Aaron Halfaker. 2017. Operationalizing conflict and cooperation between automated software agents in wikipedia: A replication and expansion of'even good bots fight'. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–33.

[60] R Stuart Geiger and David Ribes. 2010. The work of sustaining order in Wikipedia: The banning of a vandal. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 117–126.

[61] Michael Gilbert, Jonathan T Morgan, David W McDonald, and Mark Zachry. 2013. Managing complexity: Strategies for group awareness and coordinated action in Wikipedia. In *Proceedings of the 9th International Symposium on open collaboration*. 1–10.

[62] Michael Gilbert and Mark Zachry. 2015. Tool-mediated coordination of virtual teams in complex systems. In *Proceedings of the 11th International Symposium on Open Collaboration*. 1–8.

[63] Barney G Glaser and Anselm L Strauss. 2017. *Discovery of grounded theory: Strategies for qualitative research*. Routledge.

[64] Google. 2023. Search term: Wikipedia research. https://trends.google.com/trends/explore?date=all&geo=US&q=wikipedia%20research

[65] Adam M Grant. 2007. Relational job design and the motivation to make a prosocial difference. *Academy of management review* 32, 2 (2007), 393–417.

[66] J Richard Hackman and Greg R Oldham. 1976. Motivation through the design of work: Test of a theory. *Organizational behavior and human performance* 16, 2 (1976), 250–279.

[67] Scott A Hale. 2014. Okinawa in Japanese and English wikipedia. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*. 927–932.

[68] Aaron Halfaker. 2013. Wikimedia Foundation Growth Team Experiments: Onboarding New Wikipedians. https://meta.wikimedia.org/wiki/Research:Onboarding_new_Wikipedians

[69] Aaron Halfaker. 2017. Interpolating quality dynamics in Wikipedia and demonstrating the Keilana effect. In *Proceedings of the 13th International Symposium on Open Collaboration*. 1–9.

[70] Aaron Halfaker, R Stuart Geiger, Jonathan T Morgan, and John Riedl. 2013. The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline. *American Behavioral Scientist* 57, 5 (2013), 664–688.

[71] Aaron Halfaker, Aniket Kittur, Robert Kraut, and John Riedl. 2009. A jury of your peers: quality, experience and ownership in Wikipedia. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*. 1–10.

[72] Aaron Halfaker, Aniket Kittur, and John Riedl. 2011. Don't bite the newbies: how reverts affect the quantity and quality of Wikipedia work. In *Proceedings of the 7th international symposium on wikis and open collaboration*. 163–172.

[73] Aaron Halfaker and John Riedl. 2012. Bots and cyborgs: Wikipedia's immune system. *Computer* 45, 03 (2012), 79–82.

[74] Michael T Hannan and John Freeman. 1977. The population ecology of organizations. *American journal of sociology* 82, 5 (1977), 929–964.

[75] Eszter Hargittai and Aaron Shaw. 2015. Mind the skills gap: the role of Internet know-how and gender in differentiated contributions to Wikipedia. *Information, communication & society* 18, 4 (2015), 424–442.

[76] Ari Hautasaari and Toru Ishida. 2012. Analysis of discussion contributions in translated Wikipedia articles. In *Proceedings of the 4th international conference on Intercultural Collaboration*. 57–66.

[77] James M Heilman and Andrew G West. 2015. Wikipedia and medicine: quantifying readership, editors, and the significance of natural language. *Journal of medical Internet research* 17, 3 (2015), e62.

[78] Benjamin Mako Hill and Aaron Shaw. 2013. The Wikipedia gender gap revisited: Characterizing survey response bias with propensity score estimation. *PloS one* 8, 6 (2013), e65782.

[79] Benjamin Mako Hill and Aaron Shaw. 2015. Page protection: another missing dimension of wikipedia research. In *Proceedings of the 11th International Symposium on Open Collaboration*. 1–4.

[80] Kasper Hornbæk, Aske Mottelson, Jarrod Knibbe, and Daniel Vogel. 2019. What do we mean by "interaction"? An analysis of 35 years of CHI. *ACM Transactions on Computer-Human Interaction (TOCHI)* 26, 4 (2019), 1–30.

[81] Meiqun Hu, Ee-Peng Lim, Aixin Sun, Hady Wirawan Lauw, and Ba-Quy Vuong. 2007. Measuring article quality in wikipedia: models and evaluation. In *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*. 243–252.

[82] Daniel R Ilgen, John R Hollenbeck, Michael Johnson, and Dustin Jundt. 2005. Teams in organizations: From input-process-output models to IMOI models. *Annu. Rev. Psychol.* 56 (2005), 517–543.

[83] Jane Im, Amy X Zhang, Christopher J Schilling, and David Karger. 2018. Deliberation and resolution on wikipedia: A case study of requests for comments. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–24.

[84] Michał Jankowski-Lorek, Radosław Nielek, Adam Wierzbicki, and Kazimierz Zieliński. 2014. Predicting controversy of Wikipedia articles using the article feedback tool. In *Proceedings of the 2014 international conference on social computing*. 1–7.

[85] Adam Jatowt and Katsumi Tanaka. 2012. Is Wikipedia too difficult? comparative analysis of readability of Wikipedia, simple Wikipedia and Britannica. In *Proceedings of the 21st ACM international conference on Information and knowledge management*. 2607–2610.

[86] Sara Javanmardi and Cristina Lopes. 2010. Statistical measure of quality in Wikipedia. In *Proceedings of the First Workshop on Social Media Analytics*. 132–138.

[87] Karen A Jehn. 1997. A qualitative analysis of conflict types and dimensions in organizational groups. *Administrative science quarterly* (1997), 530–557.

[88] Gareth R Jones. 1986. Socialization tactics, self-efficacy, and newcomers' adjustments to organizations. *Academy of Management journal* 29, 2 (1986), 262–279.

[89] Elisabeth Joyce, Brian Butler, and Jacqueline Pike. 2011. Handling flammable materials: Wikipedia biographies of living persons as contentious objects. In *Proceedings of the 2011 iConference*. 25–32.

[90] Elisabeth W Joyce, Jacqueline C Pike, and Brian Butler. 2013. Keeping eyes on the prize: Officially sanctioned rule breaking in mass collaboration systems. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1081–1092.

[91] Nicolas Jullien. 2012. What We Know About Wikipedia: A Review of the Literature Analyzing the Project (s). (2012).

[92] Lucie-Aimée Kaffee, Kemele M Endris, and Elena Simperl. 2019. When humans and machines collaborate: cross-lingual label editing in wikidata. In *Proceedings of the 15th International Symposium on Open Collaboration*. 1–9.

[93] Andreas Kaltenbrunner and David Laniado. 2012. There is no deadline: time evolution of Wikipedia discussions. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. 1–10.

[94] Arun Kalyanasundaram, Wei Wei, Kathleen M Carley, and James D Herbsleb. 2015. An agent-based model of edit wars in wikipedia: How and when is consensus reached. In *2015 Winter Simulation Conference (WSC)*. IEEE, 276–287.

[95] Gerald C Kane, Jeremiah Johnson, and Ann Majchrzak. 2014. Emergent life cycle: The tension between knowledge change and knowledge retention in open online coproduction communities. *Management Science* 60, 12 (2014), 3026–3048.

[96] Gerald C Kane and Sam Ransbotham. 2016. Research note—content and collaboration: an affiliation network approach to information quality in online peer production communities. *Information Systems Research* 27, 2 (2016), 424–439.

[97] Brian Keegan, Darren Gergle, and Noshir Contractor. 2011. Hot off the wiki: dynamics, practices, and structures in Wikipedia's coverage of the Tōhoku catastrophes. In *Proceedings of the 7th international symposium on Wikis and open collaboration*. 105–113.

[98] Brian Keegan, Darren Gergle, and Noshir Contractor. 2012. Do editors or articles drive collaboration? Multilevel statistical network analysis of Wikipedia coauthorship. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 427–436.

[99] Brian C Keegan and Jed R Brubaker. 2015. 'Is' to'Was' Coordination and Commemoration in Posthumous Activity on Wikipedia Biographies. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 533–546.

[100] Jinyoung Kim. 2013. Wikipedians from Mars: Female students' perceptions toward wikipedia. *Proceedings of the American Society for Information Science and Technology* 50, 1 (2013), 1–4.

[101] Aniket Kittur, Ed Chi, Bryan A Pendleton, Bongwon Suh, and Todd Mytkowicz. 2007. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. *World wide web* 1, 2 (2007), 19.

[102] Aniket Kittur and Robert E Kraut. 2008. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 37–46.

[103] Aniket Kittur, Bryan Pendleton, and Robert E Kraut. 2009. Herding the cats: the influence of groups in coordinating peer production. In *Proceedings of the 5th international Symposium on Wikis and Open Collaboration*. 1–9.

[104] Aniket Kittur, Bongwon Suh, Bryan A Pendleton, and Ed H Chi. 2007. He says, she says: conflict and coordination in Wikipedia. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 453–462.

[105] Maximilian Klein, Thomas Maillart, and John Chuang. 2015. The virtuous circle of Wikipedia: Recursive measures of collaboration structures. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 1106–1115.

[106] Michael D Kocolowski. 2010. Shared leadership: Is it time for a change. *Emerging leadership journeys* 3, 1 (2010), 22–32.

[107] Travis Kriplean, Ivan Beschastnikh, and David W McDonald. 2008. Articulations of wikiwork: uncovering valued work in wikipedia through barnstars. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. 47–56.

[108] Travis Kriplean, Ivan Beschastnikh, David W McDonald, and Scott A Golder. 2007. Community, consensus, coercion, control: cs* w or how policy mediates mass participation. In *Proceedings of the 2007 ACM International Conference on Supporting Group Work*. 167–176.

[109] Stacey Kuznetsov. 2006. Motivations of contributors to Wikipedia. *ACM SIGCAS computers and society* 36, 2 (2006), 1–es.

[110] Shyong K Lam, Jawed Karim, and John Riedl. 2010. The effects of group composition on decision quality in a social production community. In *Proceedings of the 2010 ACM International Conference on Supporting Group Work*. 55–64.

[111] Shyong (Tony) K Lam, Anuradha Uduwage, Zhenhua Dong, Shilad Sen, David R Musicant, Loren Terveen, and John Riedl. 2011. WP: clubhouse? An exploration of Wikipedia's gender imbalance. In *Proceedings of the 7th international symposium on Wikis and open collaboration*. 1–10.

[112] Weiwen Leung, Haiyi Zhu, and Joseph A Konstan. 2017. The effect of emotional cues from the nfl on wikipedia contributions. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (2017), 1–21.

[113] John M Levine, Richard L Moreland, and Hoon-Seok Choi. 2001. Group socialization and newcomer innovation. *Blackwell handbook of social psychology: Group processes* 3 (2001), 86–106.

[114] Ang Li, Zheng Yao, Diyi Yang, Chinmay Kulkarni, Rosta Farzan, and Robert E Kraut. 2020. Successful online socialization: lessons from the Wikipedia education program. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–24.

[115] Jun Liu and Sudha Ram. 2011. Who does what: Collaboration patterns in the Wikipedia and their impact on article quality. *ACM Transactions on Management Information Systems (TMIS)* 2, 2 (2011), 1–23.

[116] Thomas W Malone and Kevin Crowston. 1994. The interdisciplinary study of coordination. *ACM Computing Surveys (CSUR)* 26, 1 (1994), 87–119.

[117] Owen S Martin. 2011. A Wikipedia literature review. *arXiv preprint arXiv:1110.5863* (2011).

[118] John Mathieu, M Travis Maynard, Tammy Rapp, and Lucy Gilson. 2008. Team effectiveness 1997-2007: A review of recent advancements and a glimpse into the future. *Journal of management* 34, 3 (2008), 410–476.

[119] John E Mathieu, Tonia S Heffner, Gerald F Goodwin, Eduardo Salas, and Janis A Cannon-Bowers. 2000. The influence of shared mental models on team process and performance. *Journal of applied psychology* 85, 2 (2000), 273.

[120] Joseph Edward McGrath. 1984. *Groups: Interaction and performance.* Vol. 14. Prentice-Hall Englewood Cliffs, NJ.

[121] Olena Medelyan, David Milne, Catherine Legg, and Ian H Witten. 2009. Mining meaning from Wikipedia. *International Journal of Human-Computer Studies* 67, 9 (2009), 716–754.

[122] Mohamad Mehdi, Chitu Okoli, Mostafa Mesgari, Finn Årup Nielsen, and Arto Lanamäki. 2017. Excavating the mother lode of human-generated text: A systematic review of research that uses the wikipedia corpus. *Information Processing & Management* 53, 2 (2017), 505–529.

[123] Amanda Menking and Ingrid Erickson. 2015. The heart work of Wikipedia: Gendered, emotional labor in the world's largest online encyclopedia. In *Proceedings of the 33rd annual ACM conference on human factors in computing systems*. 207–210.

[124] Mostafa Mesgari, Chitu Okoli, Mohamad Mehdi, Finn Årup Nielsen, and Arto Lanamäki. 2015. "The sum of all human knowledge": A systematic review of scholarly research on the content of W ikipedia. *Journal of the Association for Information Science and Technology* 66, 2 (2015), 219–245.

[125] David Moher, Alessandro Liberati, Jennifer Tetzlaff, Douglas G Altman, and the PRISMA Group*. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of internal medicine* 151, 4 (2009), 264–269.

[126] Jonathan T Morgan, Siko Bouterse, Heather Walls, and Sarah Stierch. 2013. Tea and sympathy: crafting positive new user experiences on wikipedia. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 839–848.

[127] Jonathan T Morgan, Katie Derthick, Toni Ferro, Elly Searle, Mark Zachry, and Travis Kriplean. 2009. Formalization and community investment in wikipedia's regulating texts: the role of essays. In *Proceedings of the 27th ACM international conference on Design of communication*. 315–316.

[128] Jonathan T Morgan, Michael Gilbert, David W McDonald, and Mark Zachry. 2013. Project talk: Coordination work and group membership in WikiProjects. In *Proceedings of the 9th International Symposium on Open Collaboration*. 1–10.

[129] Jonathan T Morgan, Michael Gilbert, David W McDonald, and Mark Zachry. 2014. Editing beyond articles: diversity & dynamics of teamwork in open collaborations. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 550–563.

[130] Jonathan T Morgan, Michael Gilbert, Mark Zachry, and David McDonald. 2013. A content analysis of Wikiproject discussions: Toward a typology of coordination language used by virtual teams. In *Proceedings of the 2013 conference on Computer supported cooperative work companion*. 231–234.

[131] Jonathan T Morgan and Aaron Halfaker. 2018. Evaluating the impact of the Wikipedia Teahouse on newcomer socialization and retention. In *Proceedings of the 14th International Symposium on Open Collaboration*. 1–7.

[132] Jonathan T Morgan, Robert M Mason, and Karine Nahon. 2011. Lifting the veil: The expression of values in online communities. In *Proceedings of the 2011 iConference*. 8–15.

[133] Jonathan T Morgan and Mark Zachry. 2010. Negotiating with angry mastodons: the wikipedia policy environment as genre ecology. In *Proceedings of the 2010 ACM International Conference on Supporting Group Work*. 165–168.

[134] Claudia Müller-Birn, Leonhard Dobusch, and James D Herbsleb. 2013. Work-to-rule: the emergence of algorithmic governance in Wikipedia. In *Proceedings of the 6th International Conference on Communities and Technologies*. 80–89.

[135] David R Musicant, Yuqing Ren, James A Johnson, and John Riedl. 2011. Mentoring in Wikipedia: a clash of cultures. In *Proceedings of the 7th international symposium on Wikis and Open Collaboration*. 173–182.

[136] Sneha Narayan, Jake Orlowitz, Jonathan Morgan, Benjamin Mako Hill, and Aaron Shaw. 2017. The Wikipedia Adventure: field evaluation of an interactive tutorial for new users. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. 1785–1799.

[137] Sneha Narayan, Jake Orlowitz, Jonathan T Morgan, and Aaron Shaw. 2015. Effects of a Wikipedia Orientation Game on New User Edits. In *Proceedings of the 18th ACM Conference Companion on Computer Supported Cooperative Work & Social Computing*. 263–266.

[138] Keiichi Nemoto, Peter Gloor, and Robert Laubacher. 2011. Social capital increases efficiency of collaboration among Wikipedia editors. In *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*. 231–240.

[139] Finn Årup Nielsen. 2012. Wikipedia research and tools: Review and comments. *Available at SSRN 2129874* (2012).

[140] Oded Nov. 2007. What motivates wikipedians? *Commun. ACM* 50, 11 (2007), 60–64.

[141] Chitu Okoli, Mohamad Mehdi, Mostafa Mesgari, Finn Årup Nielsen, and Arto Lanamäki. 2012. The people's encyclopedia under the gaze of the sages: A systematic review of scholarly research on Wikipedia. *Available at SSRN 2021326* (2012).

[142] Chitu Okoli, Mohamad Mehdi, Mostafa Mesgari, Finn Årup Nielsen, and Arto Lanamäki. 2014. Wikipedia in the eyes of its beholders: A systematic review of scholarly research on Wikipedia readers and readership. *Journal of the Association for Information Science and Technology* 65, 12 (2014), 2381–2403.

[143] Thomas A O'Neill, Natalie J Allen, and Stephanie E Hastings. 2013. Examining the "pros" and "cons" of team con flict: A team-level meta-analysis of task, relationship, and process conflict. *Human Performance* 26, 3 (2013), 236–260.

[144] Felipe Ortega, Jesus M Gonzalez-Barahona, and Gregorio Robles. 2008. On the inequality of contributions to Wikipedia. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*. IEEE, 304–304.

[145] Kim Osman. 2013. The role of conflict in determining consensus on quality in Wikipedia articles. In *Proceedings of the 9th International Symposium on Open Collaboration*. 1–6.

[146] Katherine Panciera, Aaron Halfaker, and Loren Terveen. 2009. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proceedings of the ACM 2009 international conference on Supporting group work*. 51–60.

[147] Craig L Pearce and Henry P Sims Jr. 2002. Vertical versus shared leadership as predictors of the effectiveness of change management teams: An examination of aversive, directive, transactional, transformational, and empowering leader behaviors. *Group dynamics: Theory, research, and practice* 6, 2 (2002), 172.

[148] Christian Pentzold and Sebastian Seidenglanz. 2006. Foucault@ Wiki: first steps towards a conceptual framework for the analysis of Wiki discourses. In *Proceedings of the 2006 international symposium on Wikis*. 59–68.

[149] Ulrike Pfeil, Panayiotis Zaphiris, and Chee Siang Ang. 2006. Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication* 12, 1 (2006), 88–113.

[150] Romain Picot-Clémente, Cécile Bothorel, and Nicolas Jullien. 2015. Contribution, Social networking, and the Request for Adminship process in Wikipedia. In *Proceedings of the 11th International Symposium on Open Collaboration*. 1–2.

[151] Maryana Pinchuk. 2012. Wikimedia Foundation Growth Team Experiments: Post-Edit feedback. https://meta.wikimedia.org/wiki/Research:Post-edit_feedback.

[152] Jennifer Preece and Ben Shneiderman. 2009. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS transactions on human-computer interaction* 1, 1 (2009), 13–32.

[153] Reid Priedhorsky, Jilin Chen, Shyong (Tony) K Lam, Katherine Panciera, Loren Terveen, and John Riedl. 2007. Creating, destroying, and restoring value in Wikipedia. In *Proceedings of the 2007 international ACM conference on Supporting group work*. 259–268.

[154] Yuqing Ren and Linda Argote. 2011. Transactive memory systems 1985–2010: An integrative framework of key dimensions, antecedents, and consequences. *Academy of Management Annals* 5, 1 (2011), 189–229.

[155] Michael Restivo and Arnout Van De Rijt. 2012. Experimental study of informal rewards in peer production. *PloS one* 7, 3 (2012), e34358.

[156] Joost Rietveld and Melissa A Schilling. 2021. Platform competition: A systematic and interdisciplinary review of the literature. *Journal of Management* 47, 6 (2021), 1528–1563.

[157] Richard M Ryan and Edward L Deci. 2000. Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology* 25, 1 (2000), 54–67.

[158] Jeffrey Rzeszotarski and Aniket Kittur. 2012. Learning from history: predicting reverted work at the word level in wikipedia. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*. 437–440.

[159] Jodi Schneider, Bluma S Gelley, and Aaron Halfaker. 2014. Accept, decline, postpone: How newcomer productivity is reduced in English Wikipedia by pre-publication review. In *Proceedings of the international symposium on open collaboration*. 1–10.

[160] Jodi Schneider, Alexandre Passant, and John G Breslin. 2011. Understanding and improving Wikipedia article discussion spaces. In *Proceedings of the 2011 ACM Symposium on Applied Computing*. 808–813.

[161] Jodi Schneider, Alexandre Passant, and Stefan Decker. 2012. Deletion discussions in Wikipedia: Decision factors and outcomes. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. 1–10.

[162] Jodi Schneider, Krystian Samp, Alexandre Passant, and Stefan Decker. 2013. Arguments about deletion: How experience improves the acceptability of arguments in ad-hoc online task groups. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1069–1080.

[163] Hoda Sepehri-Rad and Denilson Barbosa. 2015. Identifying controversial Wikipedia articles using editor collaboration networks. *ACM Transactions on Intelligent Systems and Technology (TIST)* 6, 1 (2015), 1–24.

[164] Similarweb. 2023. https://www.similarweb.com/top-websites/

[165] Shaun P Slattery. 2009. " edit this page" the socio-technological infrastructure of a wikipedia article. In *Proceedings of the 27th ACM international conference on Design of communication*. 289–296.

[166] David M Sluss, Rolf Van Dick, and Bryant S Thompson. 2011. Role theory in organizations: A relational perspective. In *APA handbook of industrial and organizational psychology, Vol 1: Building and developing the organization*. American Psychological Association, 505–534.

[167] C Estelle Smith, Bowen Yu, Anjali Srivastava, Aaron Halfaker, Loren Terveen, and Haiyi Zhu. 2020. Keeping community in the loop: Understanding wikipedia stakeholder values for machine learning-based systems. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–14.

[168] Thomas Steiner. 2014. Bots vs. wikipedians, anons vs. logged-ins (redux) a global study of edit activity on wikipedia and wikidata. In *Proceedings of The International Symposium on Open Collaboration*. 1–7.

[169] Bongwon Suh, Gregorio Convertino, Ed H Chi, and Peter Pirolli. 2009. The singularity is not near: slowing growth of Wikipedia. In *Proceedings of the 5th international symposium on wikis and open collaboration*. 1–10.

[170] Róbert Sumi, Taha Yasseri, et al. 2011. Edit wars in Wikipedia. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*. IEEE, 724–727.

[171] Yu Suzuki and Satoshi Nakamura. 2016. Assessing the quality of Wikipedia editors through crowdsourcing. In *Proceedings of the 25th International Conference Companion on World Wide Web*. 1001–1006.

[172] Jennifer Thom-Santelli, Dan R Cosley, and Geri Gay. 2009. What's mine is mine: territoriality in collaborative authoring. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 1481–1484.

[173] W Ben Towne, Aniket Kittur, Peter Kinnaird, and James Herbsleb. 2013. Your process is showing: controversy management and perceived quality in Wikipedia. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 1059–1068.

[174] Jesús Tramullas, Piedad Garrido-Picazo, and Ana I Sánchez-Casabón. 2016. Research on Wikipedia Vandalism: a brief literature review. In *Proceedings of the 4th Spanish Conference on Information Retrieval*. 1–4.

[175] Milena Tsvetkova, Ruth García-Gavilanes, Luciano Floridi, and Taha Yasseri. 2017. Even good bots fight: The case of Wikipedia. *PloS one* 12, 2 (2017), e0171774.

[176] Marlon Twyman, Brian C Keegan, and Aaron Shaw. 2017. Black Lives Matter in Wikipedia: Collective memory and collaboration around online social movements. In *Proceedings of the 2017 acm conference on computer supported cooperative work and social computing*. 1400–1412.

[177] Fernanda B Viégas, Martin Wattenberg, Jesse Kriss, and Frank Van Ham. 2007. Talk before you type: Coordination in Wikipedia. In *2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*. IEEE, 78–78.

[178] Ba-Quy Vuong, Ee-Peng Lim, Aixin Sun, Minh-Tam Le, Hady Wirawan Lauw, and Kuiyu Chang. 2008. On ranking controversies in wikipedia: models and evaluation. In *Proceedings of the 2008 international conference on Web search and data mining*. 171–182.

[179] Claudia Wagner, David Garcia, Mohsen Jadidi, and Markus Strohmaier. 2015. It's a man's Wikipedia? Assessing gender inequality in an online encyclopedia. In *Proceedings of the international AAAI conference on web and social media*, Vol. 9. 454–463.

[180] Steven Walling and Dario Taraborelli. 2012. Wikimedia Foundation Growth Team Experiments: Account Creation UX. https://meta.wikimedia.org/wiki/Research:Account_creation_UX

[181] Xiaoqing Wang, Brian S Butler, and Yuqing Ren. 2013. The impact of membership overlap on growth: An ecological competition view of online groups. *Organization Science* 24, 2 (2013), 414–431.

[182] Morten Warncke-Wang, Vladislav R Ayukaev, Brent Hecht, and Loren G Terveen. 2015. The success and failure of quality improvement projects in peer production communities. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. 743–756.

[183] Morten Warncke-Wang, Dan Cosley, and John Riedl. 2013. Tell me more: an actionable quality model for Wikipedia. In *Proceedings of the 9th International Symposium on Open Collaboration*. 1–10.

[184] Howard T Welser, Dan Cosley, Gueorgi Kossinets, Austin Lin, Fedor Dokshin, Geri Gay, and Marc Smith. 2011. Finding social roles in Wikipedia. In *Proceedings of the 2011 iConference*. 122–129.

[185] Robert West, Ingmar Weber, and Carlos Castillo. 2012. Drawing a data-driven portrait of Wikipedia editors. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. 1–10.

[186] Wikipedia. 2023. Editor Survey. https://meta.wikimedia.org/wiki/Editor_Survey

[187] Wikipedia. 2023. English Wikipedia. https://en.wikipedia.org/wiki/English_Wikipedia

[188] Wikipedia. 2023. Ignore all rules. https://en.wikipedia.org/wiki/Ignore_all_rules

[189] Wikipedia. 2023. Wikipedia is not paper. https://meta.wikimedia.org/wiki/Wikipedia_is_not_paper

[190] Wikipedia. 2023. Wikipedia:Article Feedback Tool. https://en.wikipedia.org/wiki/Wikipedia:Article_Feedback_Tool

[191] Wikipedia. 2023. Wikipedia:Articles for creation. https://en.wikipedia.org/wiki/Wikipedia:Articles_for_creation

[192] Wikipedia. 2023. Wikipedia:Be bold. https://en.wikipedia.org/wiki/Wikipedia:Be_bold

[193] Wikipedia. 2023. Wikipedia:Consensus. https://en.wikipedia.org/wiki/Wikipedia:Consensus

[194] Wikipedia. 2023. Wikipedia:Content assessment. https://en.wikipedia.org/wiki/Wikipedia:Content_assessment

[195] Wikipedia. 2023. Wikipedia:Deletion policy. https://en.wikipedia.org/wiki/Wikipedia:Deletion_policy

[196] Wikipedia. 2023. Wikipedia:Guide to requests for adminship. https://en.wikipedia.org/wiki/Wikipedia:Guide_to_requests_for_adminship

[197] Wikipedia. 2023. Wikipedia:Namespace. https://en.wikipedia.org/wiki/Wikipedia:Namespace

[198] Wikipedia. 2023. Wikipedia:Neutral point of view. https://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view

[199] Wikipedia. 2023. Wikipedia:No original research. https://en.wikipedia.org/wiki/Wikipedia:No_original_research

[200] Wikipedia. 2023. Wikipedia:Removing administrator rights. https://en.wikipedia.org/wiki/Wikipedia:Removing_administrator_rights

[201] Wikipedia. 2023. Wikipedia:Requests for adminship. https://en.wikipedia.org/wiki/Wikipedia:Requests_for_adminship

[202] Wikipedia. 2023. Wikipedia:Third opinion. https://en.wikipedia.org/wiki/Wikipedia:Third_opinion

[203] Wikipedia. 2023. Wikipedia:User access levels. https://en.wikipedia.org/wiki/Wikipedia:User_access_levels

[204] Wikipedia. 2023. Wikipedia:What Wikipedia is not. https://en.wikipedia.org/wiki/Wikipedia:What_Wikipedia_is_not

[205] Wikipedia. 2023. Wikipedia:Why Wikipedia is not so great. https://en.wikipedia.org/wiki/Wikipedia:Why_Wikipedia_is_not_so_great

[206] Wikipedia. 2023. Wikipedia:Why Wikipedia is so great. https://en.wikipedia.org/wiki/Wikipedia:Why_Wikipedia_is_so_great

[207] Dennis M Wilkinson and Bernardo A Huberman. 2007. Assessing the value of coooperation in wikipedia. *arXiv preprint cs/0702140* (2007).

[208] Jodi L Wilson. 2013. Proceed with extreme caution: Citation to wikipedia in light of contributor demographics and content policies. *Vand. J. Ent. & Tech. L.* 16 (2013), 857.

[209] Thomas Wöhner and Ralf Peters. 2009. Assessing the quality of Wikipedia articles with lifecycle based metrics. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration*. 1–10.

[210] Anita Williams Woolley, Christopher F Chabris, Alex Pentland, Nada Hashmi, and Thomas W Malone. 2010. Evidence for a collective intelligence factor in the performance of human groups. *science* 330, 6004 (2010), 686–688.

[211] Diyi Yang, Aaron Halfaker, Robert Kraut, and Eduard Hovy. 2016. Who did what: Editor role identification in Wikipedia. In *Proceedings of the International AAAI Conference on Web and Social Media*, Vol. 10. 446–455.

[212] Taha Yasseri, Robert Sumi, András Rung, András Kornai, and János Kertész. 2012. Dynamics of conflicts in Wikipedia. *PloS one* 7, 6 (2012), e38869.

[213] Bowen Yu, Yuqing Ren, Loren Terveen, and Haiyi Zhu. 2017. Predicting member productivity and withdrawal from pre-joining attachments in online production groups. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. 1775–1784.

[214] Shumin Zhai. 2014. TOCHI turns twenty. , 4 pages.

[215] Dell Zhang, Karl Prior, and Mark Levene. 2012. How long do Wikipedia editors keep active?. In *Proceedings of the Eighth Annual International Symposium on Wikis and Open Collaboration*. 1–4.

[216] Lei Zheng, Christopher M Albano, Neev M Vora, Feng Mai, and Jeffrey V Nickerson. 2019. The roles bots play in Wikipedia. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–20.

[217] Haiyi Zhu, Robert Kraut, and Aniket Kittur. 2012. Effectiveness of shared leadership in online communities. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 407–416.

[218] Haiyi Zhu, Robert E Kraut, Yi-Chia Wang, and Aniket Kittur. 2011. Identifying shared leadership in Wikipedia. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 3431–3434.

[219] Haiyi Zhu, Bowen Yu, Aaron Halfaker, and Loren Terveen. 2018. Value-sensitive algorithm design: Method, case study, and lessons. *Proceedings of the ACM on human-computer interaction* 2, CSCW (2018), 1–23.

[220] Haiyi Zhu, Amy Zhang, Jiping He, Robert E Kraut, and Aniket Kittur. 2013. Effects of peer feedback on contribution: a field experiment in Wikipedia. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2253–2262.

[221] Jinlong Zhu, Zhenyu Liao, Kai Chi Yam, and Russell E Johnson. 2018. Shared leadership: A state-of-the-art review and future research agenda. *Journal of Organizational Behavior* 39, 7 (2018), 834–852.