Momentum17: Vielfalt

Who Edits? - A Literature Review of Wikipedia’s Editor Community and Governance

Alexandra Price, Leopold-Franzens-Universität Innsbruck
Univ.-Prof. Dr. Leonhard Dobusch, Leopold-Franzens-Universität Innsbruck

Introduction

By declaring itself as the “free encyclopedia that anyone can edit”, Wikipedia aims to create a freely shared platform representing all human knowledge. The community-created encyclopedic giant has management to establish itself as one of the most influential sources of online information worldwide and has become indispensable for all kinds of learning and educational purposes. Theoretically, Wikipedia should represent an unbiased and neutral source of knowledge; its self-proclaimed guidelines include notability, verifiability, as well as neutral point of view (Wagner et al., 2015; Wagner et al., 2016). Yet, in practice, Wikipedia’s community is not diverse and a survey conducted by the Wikimedia Foundation (WMF), the nonprofit organization behind the various Wikipedia projects, revealed that the percentage of female contributors barely reaches 10 percent. Although Hill and Shaw (2013) corrected this statistic by estimating the actual female contribution to be slightly higher at 16.1 percent, Wikipedia has still failed to reach a minimal gender balance (Massa, & Zelenkauskaitė, 2014). In addition, Wikipedia is facing a steady decline in the number of active contributors and recent research suggests that this primarily can be traced back to the failure to retain newcomers (Haflacker et al., 2013).

The reason why Wikipedia editors’ diversity is of importance is related to the fact that it is one of the most accessed online sources of knowledge, containing more than 5 Million articles available in over 280 language editions (Wikipedia Homepage). Because so many people use the free encyclopedia, it is crucial to understand that what these pages reflect is predominantly the point of view of a white male population. The issue of gender distribution has also been tackled by the WMF itself (Antin et al., 2011; Massa, & Zelenkauskaitė, 2014) and one of the central goals of their 2010 five-year strategic plan was “doubling the percentage of female editors to 25 percent by 2015” (Wikimedia Foundation, 2011, p.6). However, recent studies show that they have failed to reach this goal. According to MacAulay and Visser (2016) the corporatized diversity initiative designed by the WMF is
more effective when it comes to organizational growth and its failure can be explained by insufficiently addressing the underlying reasons behind the gender gap.

In light of the above described situation, various studies have tried to analyze and understand the reasons as well as extent of the pronounced gender gap in Wikipedia (MacAulay, & Visser, 2016). We suggest that the gender gap might be related to the governance model currently in place. Hence, this paper aims at gaining a better understanding of what might be behind Wikipedia’s gender gap by means of a literature review. The paper starts with reviewing the literature regarding diversity with a focus on community, followed by governance literature with a focus on diversity. Before coming to the discussion we also discuss existing literature and findings on the topic of Wikipedia and diversity.

**Diversity with a Focus on Community**

“Homogeneous societies are quickly becoming a relic of the past” (Essed, 1996, p.1). Intensified migrations, intercultural communications, cross-cultural coalitions, transnational cooperation’s, and international agreements all confirm the fact that we are now in the age of diversity. However, life and work in such a multi-diverse society comprises conflicts, contradictions, and uncertainties around the experience of ethnic or national identity. The challenge we are facing is how we can create justice and respect between individuals and groups of people at a time when worldwide migrations intensify and new community situations continue to emerge (Essed, 1996).

The topic of diversity has gained consistent and increased attention by the business media, organizations, the popular press, as well as researchers (Roberson, Ryan, & Ragins, 2017). Despite the large body of knowledge and literature on diversity, many researchers agree it is “confusing – difficult to understand and difficult to synthesize” (Harrison & Klein, 2007, p. 1200). This can partly be explained by the fact that there have not been consistent findings and cumulative insights, but also because the literature on diversity itself is very diverse. Much of the research investigates differences in demographic variables, such as gender, ethnicity and race, tenure, and education, whereas other research addresses non-demographic variables, such as values, attitudes, affect, conscientiousness, network ties. In addition, the wide-ranging theoretical perspectives used in diversity research also make synthesis difficult. The different perspectives often propose contradictory effects (Harrison & Klein, 2007). Despite this, there are a few concepts, some borrowed from organizational diversity, that stand out and hence should be discussed.
Much of the diversity research is strongly related to social-psychological theories of cross-group relations. Especially self-categorization (Tajfel, 1978) and social identity theories (Turner, 1982), which propose practises through which individuals comprehend their social environment and locate themselves within it, help to explain the process through which individuals identify with others through their group membership. The theories argue that individuals’ self-definitions are influenced by their group membership and, hence, are interested in enhancing their self-concept by trying to create a positively valued distinctiveness for their groups. Consequently, they participate in social comparison to distinguish between so called in-groups and out-groups, which emphasise similarities members of the same group and differences among individuals of different identity groups. In addition, self-categorization theory argues that individuals also use demographic characteristics to classify social categories and use this as the basis for defining themselves as part of a social group (Turner, 1982). Likewise, the similarity-attraction paradigm (Byrne, 1971) articulates that individuals are attracted to others who possess similar characteristics, which effects their social interaction and intergroup relations (Roberson, Ryan, & Ragins, 2017).

The above described theories provide the key concept for relational demography (Tsui, Egan, & O’Reilly, 1992; Tsui & O’Reilly, 1989), which argues that the amount of demographic likeness within work units influences individuals’ attitudes and behaviours. Moreover, based on the notion that demographic attributes are important features of how individuals define themselves, relational demography theory argues that work units will automatically become more attractive because others in the work unit share such attributes. As an individual is driven to maintain a positive self-evaluation, an increased demographic similarity within a group will create better work relations and positive attitudes. However, since demographic attributes form the basis for intergroup differentiation, differences on crucial dimensions of identity can possibly impair social processes within the work unit, e.g. communication and cohesion (Roberson, Ryan, & Ragins, 2017).

However, contrary to the above mentioned identity-related diversity perspectives, research has also shown that diversity can have positive informational benefits. This perspective, either cited as the informational/decision-making perspective (Williams & O’Reilly, 1998), or the value-in-diversity hypothesis (Cox & Blake, 1991), argues that differences in groups enhances the range of perspectives, as well as other cognitive resources. In addition, research argues that such dissimilarities expose individuals to minority options and other creative ideas and solutions, while also giving access to larger and more diverse
social networks. Therefore, with greater access to relevant knowledge and information, groups have a greater problem solving and decision making ability. In other words, this viewpoint argues that diversity as an informational source can create performance benefits (Roberson, Ryan, & Ragins, 2017).

Apart from the value-in-diversity theory, diversity research has also drawn from social information processing theory. This perspective reasons that social interactions are essentially organizational events that create shared meaning and hence can be used to resolve ambiguity and procedure impressions of individuals and groups (Weick, 1995). Similarly to social-psychological identity construction and maintenance theory, individuals use such interactions to define themselves and others, and form a sense of their work environment. In addition, as argued by climate theorists, individuals gather information from their work environment and other relevant stimuli to interpret, describe, as well as attach meaning to their workplace (Schneider & Reichers, 1983). Thus, demographic differences and other psychologically salient events involving diversity are used for sensemaking and forming perceptions about their work environments and others with whom they work. Put simply, diversity and related experience can influence diversity climate perceptions (Roberson, Ryan, & Ragins, 2017).

Having discussed relevant perceptions of diversity, one might pose the question how diversity relates to community. Although many researchers agree that a sense of community is a desirable component and positive outcome of social functioning, one of the main points of critique is its focus on uniformity, regularity, and balance in community life (Townley et al., 2011; Wiesenfeld, 1996). Wiesenfeld (1996) articulates this matter as follows: “Community refers to a homogeneous group of individuals, clearly distinguishable from others” (p.337). Similarly, other researchers have defined the psychological sense of community as the perception of resemblance to others with a high degree of homogeneity and uniformity in terms of group values and norms. In particular, the component of membership in a community makes it clear that communities form boundaries of who belongs and who does not belong. Furthermore, a common community development dynamic is that individuals seek similar others (Townley et al., 2011). As pointed out by McMillan (1996): “If one can find people with similar ways of looking, feeling, thinking, and being, then it is assumed that one has found a place where one can safely be oneself” (p.321). Hence, diversity in terms of heterogeneity and community homogeneity involve processes and relationships among individuals and groups that are challenging to reconcile and, hence, many communities might not be diverse (Townley et al., 2011).
Governance with Focus on Diversity

In the traditional sense, the term governance is commonly used in the private, as well as the public sector. It comprises formal and informal structures, norms, and practices, global and local arrangements, and spontaneous and deliberate systems of control (Williamson, 1998). In terms of the public sector, governance can be defined as the institutions and processes that define how power is implemented, how individuals are given a voice, as well as how decisions are made based on public concern. Alternatively, governance in the private-sector is usually connected to the firm and direction of the individual organization or to institutions that sustain the stability of markets. Therefore, governance may either refer to controlling and leading multinational operations, or to a system of rules which warrants freedom of, access to information, entry into the market, and the promise of contracts (Burnell, 2000).

In recent years, capital market participants are increasingly focusing their attention on the governance of organizations in reaching the social- and financial goals of their stakeholders. With this in mind, the new tendency is to increase diversity in governance. For example, by introducing the requirement that a major part of the board- and audit committee members are to be independent, a greater diversity of interests and opinions is to be introduced into corporate governance ensuring that management’s control over decision processes is kept within bounds (Francoeur, Labelle, & Sinclair-Desgagné, 2008). Research has also indicated that incorporating female board members can enhance the financial performance of an organization, as well as the quality of the monitoring role of the organization’s board (Gul, Srinidhi, & Ng, 2011).

Commons-based peer production, e.g. Wikipedia and open source software, are frequently defined as “self-organized, resource governance regimes” (Müller-Birn, Dobusch, & Herbsleb, 2013, pp. 81). The reason for contributing to these different types of communities is not related to monetary incentives. Learning and other intrinsic motivation are the hidden agenda in this context. Hence, an open source community (OSC) has to consider that incentive structures, rules, and motivation enhance the coordination of contributions, as well as the involvement contributors and therefore can inspire individuals to create contributions. The traditional methods, which focus on control, supervision and monitoring of the internal processes, as defined above, often neglect this point. Although the benefits of increasing diversity might be the same, these types of government instruments are often not compatible with the micro and macro configuration of OSC and the way they operate and, hence, might not work (Latteman & Stieglitz, 2005).
According to O’Mahony (2007) a community managed governance model is comprised of five distinct principles, namely *independence, pluralism, representation, decentralized decision-making*, and lastly *autonomous participation*.

The first principle, independence, refers to the fact that an independent community does not rely on resources from another organization, but instead on the support of a diverse body of contributors. Control over the community does not dependent on a sponsor but is located within community itself. While a community might have a form of hierarchy, this hierarchy emerges independently from the reward and compensation structure that can be found in organizations. Whether or not a community is independent can be determined by understanding its decision-making structure, the base of material support, as well as independence from authority structures (O’Mahony, 2007).

The second principle O’Mahony (2007) mentions is that of pluralism. A community that is pluralistic allows many approaches, theories and methods to be credible and legitimate in following a course of action. Each holder of a specific perspective is able to advance their own approach. Community-manged projects that mange to ensure pluralism might prevent the formation of a dominant control group and can stimulate the creation of a multilateral member base. A pluralistic community should offer equal entrance to the code and process of the community, as well as an equal opportunity in becoming a cherished member of the community. A community’s pluralism can be ascertained by examining its demographic, geographic, and functional diversity of the members, how the members of a community deal with conflict, and how leaders develop and the degree of turnover.

The third principle, representation, is founded on the notion that because it is challenging to equally distribute control to many, especially in the case of large communities, communities often have to establish a system where the interest of their members is represented. Communities often start as democracies and eventually develop towards representative democracies in order to help them cope with their scale. The membership structure and rights of a community, the extent to which members can voice their opinion, and the accountability of those in authority can help understand the extent to which a community provides representation (O’Mahony, 2007).

The fourth principle, decentralized decision-making, is related the fact that community managed governance models distribute certain rights to their members. If only one party holds all decision rights, it is not a community-managed project anymore. Overall, decision-rights can be categorized into three different levels: code-level, sub-project level, and community wide decisions. Code-level decisions involve the determination of new
features and modifications and are to be incorporated in the code base, whereas some of these decisions might happen within a sub- or top-level project. A community-wide decision, on the other hand, involves a decision on the basis of the assets, organization, or process of the community. To understand a community’s decision-making, one can examine how members acquire code-level access- and decision-making rights, as well as the extent to which project activities and communications are publicly accessible (O’Mahony, 2007).

Lastly, the principle of autonomous participation is founded on the notion that a community will not last long enough to create a governance model if possible members cannot freely contribute on their own conditions. What partly attracts individuals to a community is the prospect of improving their skills, solving problems, and the opportunity to learn. Therefore, whether informal or formal, a governance model should allow members to contribute towards their own abilities, interests, and motivations. In order to examine to what extent a community encourages autonomous participation, one could study the rate of newcomers and socialization opportunities for newcomers, in addition to the rate of idea generation (O’Mahony, 2007).

**Wikipedia and Diversity**

The sum of all knowledge for every single human being; we can’t do that by leaving people out. Our vision is about more than providing universal access to all forms of knowledge. It’s about creating an inclusive culture. It’s about inviting others to join in and thrive with us. It’s about embracing human diversity. (Wikimedia Foundation, 2017)

Although there have been multiple efforts to increase female participation, there is still a large gender gap in Wikipedia with severe consequences (Lam et al., 2011). Research has suggested that certain legitimate viewpoints are partly not represented since the editors essentially contributing most of the content on Wikipedia are generally not mindset-wide or socio-demographically representative of society, the average Wikipedia reader, or even the average Internet user. The sparse information about who actually contributes to Wikipedia proposes that active editors only represent a small proportion of the sociodemographic scope of the offline-population. The findings of the online survey, the “UNU-Merit Wikipedia Survey”, conducted in 2008 with more than 300,000 Wikipedians, and from an independent survey conducted by Hill and Shaw (2013) showed that of all the participants only 16.1% are
female contributors, 24.5% are occasional and 9.2% average contributors (Glott, Schmidt, & Ghosh, 2010). According to Flöck, Vrandečić, & Simperl (2011) these findings prove that “if only such a narrow socio-demographic selection of Internet users – or of the worlds population as a whole – passes the threshold to authorship, the viewpoints inherently introduced to the article via their contributing will be equally representative of the general public or even the common Wikipedia readership” (p. 3). In addition, this disproportioned state of affairs is worsened in light of the fact that every contributor is biased along numerous dimensions, whether this is deliberately due to an option or agenda or unconsciously because it is impossible to know of every existing perspective of a topic. Although openness is a precondition for the diversity of Wikipedia’s editors, self-selection can nevertheless lead to some formation of homogeneity within the system (Flöck, Vrandečić, & Simperl, 2011).

According to technology adoption literature and volunteering literature research there should not be a large gender gap in Wikipedia’s editor community. In his research, Taniguchi (2006) found that females engage more in volunteer work than males. Moreover, Wilson (2000) articulates four essential reasons for this: they value helping others more, they exhibit greater altruism and empathy, they take on the gender-specific role of taking care of others, and they consider volunteering as a component of their social life. Simply put, research suggests it is more likely that females will volunteer their time to editing Wikipedia than males, although they might invest less time editing if they do not have a social connection to the editor community (Lam et al., 2011).

According to the technology adoption literature, females may be slower in adopting new technologies compared to males. Venkatesh, Morris, & Ackerman (2000) suggests that males are more likely to adopt new technologies and that they are influenced less by social media norms affiliated to a technology or supposed difficulty of a new technology. On the other hand, research in the area of gender differences related to internet and social media adoption offer more encouraging findings. According to general populace surveys conducted by the Periodic Pew Research Center, the Internet usage was skewed towards males between 2000 and 2004, but that this gap has since been vanished. Furthermore, it is now more likely for females to participate in specific social media sites and it is more likely that females will tweet and blog. Even the traditionally male dominated activity of online gaming is becoming more popular with females (Lam et al., 2011).

All in all, these findings suggest that Wikipedia should have a reasonable gender balance. Yet, multiple studies have indicated the opposite. In their surveys of university students, Lim and Kwon (2010) found that while all participants had used Wikipedia, women
used it less regularly and questioned the quality more than men. In 2009 the Wikipedia Foundation authorized a survey of Wikipedia users and the results did not only show a large gender gap imbalance among readers, i.e. 25% females and 75% males, but an even larger imbalance among editors, i.e. 13% females and 87% males (Glott, Schmidt, & Ghosh, 2010; Lam et al., 2011).

Many researcher have speculated as to why Wikipedia has such an immense gender gap and many different theories have been proposed. One of these is the reflection theory, which proposes that the gender imbalance in Wikipedia is a direct reflection of society itself. In reality the power of interpretation is connected to social status and being granting oneself authority in a certain field of knowledge does not speak for an equally distributed society. However, the reflection theory does not explain why the Wikipedia’s proportion of women is lower than other online-communities and forums (Dobusch, 2013).

The Hacker-culture theory argues that the Wikipedia’s ideology of free knowledge combined with the usage of an open copyright license are strongly correlated to the male dominated hacker-culture of free- and open-source software. In fact, the share of women in open-source software projects is indeed even lower than in Wikipedia (Dobusch, 2013).

One of the consequences of being open to everyone, is that Wikipedia is also accessible for difficult and conflict-seeking individuals. According to the troll-theory, even a minority of trolls can negatively influence the atmosphere and chase users, especially females, away. Wikipedia itself also realized the harsh conversation tone used in debates about relevance and removal to be an issue and addressed it with the project “Teahouse”. The teahouse should represent a friendly environment, where new editors are gradually introduced to Wikipedia’s culture and can ask questions and get acquainted to others (Dobusch, 2013).

In line with the hacker-culture and troll theory, the path-dependency theory argues that a low percentage of women will lead to even less female newcomers and vice-versa. The contradicting issue is that once a strong male-associated conversation tone has been established, it is difficult to revoke and change, because exactly those people who are will to drive change are put off (Dobusch, 2013).

The leisure theory articulates that online collaboration is a voluntary engagement and that women are still responsible for the majority of reproduction work and, hence, simply have less leisure time available for contributing to Wikipedia. The same principle applies to the lower contribution rate of poorer countries; even though the people have access to computers and Internet, they simply have less time available to spend contributing to Wikipedia (Dobusch, 2013).
The usability theory argues that until recently one needed a basic understanding of the wiki-syntax in order to be able to contribute to Wikipedia. In addition, the usability and design of the Wikipedia pages appear very outdated. This could partly explain both the overall loss in editors, as well as the gender gap (Dobusch, 2013).

Lastly, the bot theory is based on the notion that technically advanced Wikipedia users installed so called “bots”, which are scripts that automatically take over recurring and borings tasks such as categorizing and spellchecking. According to this theory, the bots have replaced the tasks that newcomers previously did, hence introducing an additional barrier for newcomers (Dobusch, 2013).

Discussion

In light of the above findings, there are a few important insights regarding the diversity of Wikipedia that should be discussed.

Social identity- and self-categorization theory might help explain that since Wikipedians have a strong sense group membership, they might engage in social comparison in order to distinguishing between so called in-groups and out-groups, for example between long-term editors and newcomers. In addition, the characteristics of the Wikipedians, i.e. white and male, might be an indicator that demographic characteristics, such as gender, are used to define their social group, creating distinctiveness for the Wikipedia community. The similarity-attraction paradigm further could help support this argument, since it explains why Wikipedia attracts individuals with similar characteristics.

Relational demography also might add to understanding Wikipedia’s large gender gap, since it argues that the demographic likeliness within a group can influence individual’s attitudes and behaviours, creating better relations and positive attitudes within the group. This, for example, might be an explanation as to why Wikipedians are so reluctant towards the topic of encouraging female participation, as well as why newcomers are easily chased away.

Without a doubt, the sense of community is rather strong within the Wikipedia. Combined with what we discussed in the chapter regarding Wikipedia and diversity, one can conclude that the community itself is rather uniform and homogenous, sharing similar values and norms. In addition, the informational/decision-making perspective alludes to the fact that diversity within a group can have important informational benefits. This might also help understand why previous studies have found systematic biases, e.g. gender biases, within
Wikipedia’s content. For instance, Wagner et al. (2015) found that women are portrayed differently than men are portrayed in Wikipedia.

However, how is this related to governance? When comparing O’Mahony’s (2007) community governance model to that of Wikipedia, one might question to what extent governance is related to the current gender gap. Although Wikipedia’s governance might be independent and autonomous, one could argue that it is not pluralistic and that the representation system is not neutral. According to O’Mahony, a pluralistic community allows for many approaches, theories and methods to be legitimate and individuals are able to voice their opinion and have equal opportunities, which should also be evident in their representation system. However, whether these aspects qualify for Wikipedia and whether females have the same opportunity at becoming a valued community member than males remains open.

All Wikipedia language editions share the same set of basic guidelines and policies, e.g. a neutral point of view. However, it is important to note that each language edition also represents a unique and independent case, shaped by specific characteristics, such as year of creation, cultural settings, number of active users, as well as the governance of the platform. In their article, Massa and Zelenkaukskaite (2014) found strong deviations between different language versions regarding the percentages of females among users who stated their gender, e.g. the 39.93% of the editors of the Slovenian language edition are female, whereas only 15.16% female editors in the German Wikipedia version. This might indicate that differences in governance might have an effect on the diversity of the respective Wikipedia. While self-organized governance regimes might be very affective in some ways, they also might reinforce conformity within a community.

In conclusion, we suggest reviewing the governance of different Wikipedia language versions and comparing this to Massa and Zelenkaukskaite (2014) findings regarding the deviation in female contribution across various Wikipedia platforms. This might help understand the gender gap from a different perspective and also might help to reduce it in the future.
Sources


Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. Academy of management review, 32(4), 1199-1228.


