Forms of Adaptation of Newcomers in Wikipedia as Online Sharing Knowledge Project

Anna Karczewska
Częstochowa University of Technology, Częstochowa, Poland
anna.karczewska@pcz.pl

Abstract: Wikipedia is one of the largest virtual communities of practice aimed at building and sharing knowledge online. Its ambitious goals require the cooperation of millions of users around the world. The development of the project has slowed down over the last few years, hence the acquisition and retention of new users is becoming more and more challenging. Consequently, properly organized adaptation, and encouragement of newcomers to become a part of the community emerges as a crucial issue for the existence of Wikipedia. This paper presents the results of the quantitative research carried out with the use of the content analysis method on the different language versions of Wikipedia. The issues investigated included the following aspects: the level of advancement of the newcomers’ adaptation process, the use of various adaptation tools, forms of welcoming and enhancing newcomers to learn and edit the Wikipedia content, while also their potential effect on the productivity and engagement of users. The results of the research indicate great diversity among the tools used, depending on the language version of the project. Not all the statements directed to the newcomers are encouraging and friendly, as some cultures create more restrictive and punitive content than the others. The latter leaves the newcomer with the responsibility for the process of learning the rules of behavior of the Wikipedia community. The more advanced the adaptation process is, the more active and productive users are. The conclusions of the research may be of use to different virtual communities of practice and organizations which face the problem of adapting newcomers in the online activity connected with knowledge sharing.

Keywords: adaptation, knowledge sharing, newcomers, organization, Wikipedia.

1. Introduction

Contemporary online knowledge building and sharing projects like Wikipedia resemble a rather disorganized bazaar-like structure instead of a traditional cathedral structure. In the concept of the “cathedral”, the decisions are usually taken by central management. The “bazaar” describes an open approach, where the online community is in charge of the whole system (Capiluppi, Michlmayr, 2007, p. 32). A bazaar-like world “behaves in many respects like a free market or an ecology, a collection of selfish agents attempting to maximize utility which in the process produces a self-correcting spontaneous order more elaborate and efficient than any amount of central planning could have achieved” (Raymond, 2001, p. 22). This is most probably the source of the success of Wikipedia and other online open-source projects.

Wikipedia’s content can be publicly edited by anyone, which is another reason for its rapid growth. However in the years 2007-2008, the number of users started to decline. Therefore, one of the significant subjects of research became the adaptation and retention of the new users (newcomers). Newcomers can be defined as novice participants who are just starting their activity in the project (Steinmacher et al., 2015 p. 81). The newcomers’ adjustment to the new environment has been defined as socialization, or in other words, the process of learning the behaviour and attitudes crucial to playing a role in an organization.

Newcomers in projects such as Wikipedia or open source software often face social or technical barriers that hinder learning and may discourage them from becoming a part of the project (Halfaker et al. 2013). As Choi points out (2010, p. 109), taking into consideration the high turnover rates and lack of input from the majority of users, research on the strategies that support positive relationships between online communities and their members have significant theoretical and practical implications.

The aim of this paper is to identify the forms of adaptation/socialization of the new users (newcomers) of different language versions of Wikipedia and study the relations between the level of advancement of newcomers’ adaptation process, as well as the level of activity and engagement of the users in creating the content of Wikipedia. The paper structure includes the analysis of the literature of the subject, methodology of the conducted research, results, discussion on the obtained results, while culminating in the conclusions.
2. Literature review

Retaining users who have started participating in online projects has attracted notable attention among researchers (Lai & Chen, 2014; Galehbakhtiari & Hasangholi Pouryasouri, 2015; Yan, Wang, Chen, & Zhang, 2016). The majority of literature on this subject with relation to Wikipedia originated after 2007, when the decline of the participants was observed. The ways of Wikipedia newcomers’ socialization have also been investigated over the last several years (Steinmacher, Wieses, & Gerosa 2012, Steinmacher et al. 2015, Konieczny 2018, Morgan and Halfaker 2018). Every newcomer can edit Wikipedia, albeit (s)he needs to learn basic social and technical aspects of the project before making an appropriate contribution. The “newbies” seek help in Wikipedia forums when in doubt. If they receive not polite answers, they will most probably drop out (Steinmacher, Wieses & Gerosa 2012). Although gaining new participants is crucial to the long-term survival of the community, it can be a source of disagreements in the short-term. As newcomers may more or less unintentionally behave contrary to the code of conduct and therefore impede Wikipedians’ activity, it can lower the experienced Wikiedians’ motivation to participate. Newcomers may encounter hostility from more experienced users who perceive their mistakes as a danger to the quality of Wikipedia. (Morgan & Halfaker, 2018, p. 1).

Halfaker, Geiger, & Terveen (2014) point out that one of the problematic contradictions raised in the literature of the subject is the one between the retention of Wikipedia newcomers and the content’s high quality. New quality control measures in Wikipedia, such as autonomous anti-vandal bots and specialized user interfaces designed to revert destructive edits, have been very effective in minimizing potential harm, yet also very often dissuaded newcomers. Additionally, P. Konieczny (2018, p. 157) indicates that in spite of the fact that among Wikipedia policies there are several points that encourage friendly treatment of all editors, in reality, newcomers often meet with overworked volunteers who spend a long time repairing acts of malicious vandalism. As a result, they do not often have patience to be friendly to users who make mistakes.

One of the studies also found that, among the most researched barriers to the adaptation of newcomers, the most noted are the following: the lack of social interaction with members of the project; not receiving an answer, and the newcomers’ small technical experience. What may also constitute a barrier is the fact that newcomers need to take action and search by themselves for the proper knowledge that is required in the project. (Steinmacher et al., 2015, pp. 80-83). Thus, the most evidenced barriers are related to socialization, as well as knowledge sharing.

Another part of the literature focuses on the effective tools and techniques facilitating the retention of newcomers. An important issue raised is how to solve the newcomers’ problem of finding a suitable task to start with. Encouragement from more experienced Wikipedians for a newcomer to find a relatively easy task to work on, or the tool which would automatically suggest the tasks is recommended. The next solution is to provide a mentor for a newcomer, which may increase the odds of a newcomer staying longer in the project (Canfora et al., 2012, Steinmacher, Wieses, & Gerosa, 2012). Morgan et al. indicated that user-friendly tools, positive interaction with experienced users and safe spaces for newcomers may be promising tools for effective adaptation (Morgan et al., 2013). According to their research results, newcomers invited to the “Teahouse” forum (dedicated to newcomers) were retained at a higher rate than editors who did not receive the invitation (Morgan & Halfalker, 2018, p. 1). It is also advisable to create more spaces dedicated to communication processes. As Jarvenpaa and Keating (2021, p. 10) noticed, cooperation in virtual teams fails due to the lack of trust and the way to increase trust is to create the possibility of free communication.

Choi et al. researched the impact of certain socialization tactics on the newcomers’ involvement in online projects. Seven tactics used most often have been distinguished as follows: invitations to join, offers of assistance, welcome messages, positive feedback, requests to work on tasks, constructive criticism, as well as personal-related comments. Although most new users made fewer edits over time, their declines were reduced in the case of those who have been socialized with the afore-mentioned tools. (Choi, 2010, p. 107). Likewise, P. Konieczny (2018) indicates that a major reason for burnout/dropout among volunteers is that of interpersonal conflicts, which shows that good relationships in the online community are a significant factor in users’ retention.

One of the most well-known theoretical models of socialization is the one developed by Van Maanen and Schein (1979). Referring to this typology, Jones (1986) assumed that collective, sequential, formal, fixed,
serial, and investiture tactics, so-called “institutionalized socialization”, help newcomers to adjust to their new role better, while also prevent their uncertainty. Meta-analytical articles suggest that newcomers experiencing more institutionalized socialization tactics in online projects may have more certainty about their new role, feel more accepted by the community, have a higher level of satisfaction with their activity, and show greater organizational commitment. They are also less likely to leave the project and perform their work more effectively (Saks, Uggerslev, & Fassina 2007; Choi, 2010, p. 108).

Moreover, Chao et al. (1994) claim that newcomers who possess greater knowledge about such project elements as internal group policy, the meaning of jargon or group goals and values tend to have a higher level of involvement and satisfaction with their activity. As Choi et al. indicate (2010), Wikipedia uses more “individualized socialization” techniques in the adaptation of newcomers. The project enables newcomers to edit articles with no training, even though it has numerous policies and recommendations on how users should behave. It is crucial to provide newcomers with the knowledge about community, preferably in an institutionalized way.

In the most recent research, Konieczny (2020) and Jemielniak (2019) related to the afore-mentioned subjects. They raised the issue of encouragement and retention of specialists and experts in Wikipedia in order to protect its content quality. Whereas Yan and Jian’s study (2017) confirmed that in accordance with the social exchange theory, receiving answers of high quality positively influenced the future knowledge seeking behaviour of newcomers. Social responses such as answers to the newcomers’ questions were found to have a strong positive influence on future knowledge contribution of newcomers.

The socialization and retention of newcomers in Wikipedia and other online open source projects have been studied most extensively in subject-related literature. Yet, the general overview of the adaptation process and tools in different Wikipedia language versions, especially from the cultural perspective, have not received in-depth attention. This paper attempts to fill the knowledge gap in this respect.

3. Methodology

The aim of the presented research was to identify the forms of adaptation of new users of Wikipedia. Other analysed issues were the relations between the level of advancement of the newcomers’ adaptation process in different language versions (wikis) and the level of activity and engagement of the users in creating the Wikipedia content. These aspects have been addressed in the cultural context indicated by the chosen Hofstede cultural dimensions. Chosen non-content pages were analysed, while the method of content analysis was adopted. The research was conducted in January and February 2022 and all the statistical calculations were conducted with the use of the Statistica software.

Based on the studied literature, the following hypotheses have been set:

H1: The more advanced the adaptation process is, the more active users there are in Wikipedia.

H2: The more advanced the adaptation process is, the more productive active users there are in Wikipedia.

H3: The adaptation process in Wikipedia depends on certain cultural dimensions of a society:
H3a: the more feminine society is, the more advanced the adaptation process is.
H3b: the more uncertainty society avoids, the more advanced the adaptation process is.
H3c: the more collectively oriented society is, the more advanced the adaptation process is.

In order to measure the variables in hypotheses, chosen indicators have been applied in the research.

1. The indicator of adaptation process advancement was the number of adaptation tools used in a given language version of Wikipedia.

2. The indicator of being an active user is the fact of not only registering the account but also constantly creating the content of Wikipedia.

3. The indicators of productivity were the ratio of the number of articles to the number of active users, and the number of total pages to the number of active users, as well as the number of edits to the number of active users.

4. Femininity-masculinity, uncertainty avoidance and individualism-collectivism indices have been adopted as indicators from G. Hofstede’s research (Hofstede, 2007).
The “feminine” culture of a country is characterised by social gender roles overlapping, as well as taking care of relationships and an individual’s needs. The “uncertainty avoidance” determines the extent to which the members of a society feel anxious about ambiguous situations and create institutions helping to avoid them. While the level of “individualism-collectivism” of a society is the degree of interdependence a society keeps among its members, in individualist countries people usually look after themselves and their closest family only, whereas in collectivist ones people are loyal to the “in-groups”, which take care of their members (Hofstede, 2007).

In the cases of hypotheses 1 and 2, all the language versions of Wikipedia were placed under review (325). In the case of hypothesis 3, the author researched only some language versions (15). The choice of the surveyed versions was limited by the availability of the analysed data, both the Hofstede indexes, as well as automatic Welcome Templates in Wikipedia language versions. Welcome templates are schematic messages prepared to be placed on user pages by any user. Whereas automatic welcome templates are schematic messages placed automatically on the user pages of new registered accounts by bots or scripts. The Village Pump is the public forum for all the users in a given language version. Meanwhile, Teahouse is a parallel public forum dedicated primarily to newcomers. Welcoming committees consist of members welcoming newcomers who have already made constructive edits and help them get started.

The first step of the actual research process started with counting the tools that occur in the adaptation process of newcomers in Wikipedia. The counting was conducted based on interwiki links, which are links leading to the analogical pages in different wikis. In the second step, each language version was checked in terms of the presence of an automatic welcome message for new users. Then, the chosen statistical data characterizing wikis (e.g. number of articles, number of edits, total number of pages) were obtained from the non-content page: https://meta.wikimedia.org/wiki/List_of_Wikipedias. In the third step, Spearman’s rho correlations between chosen variables characterizing wikis and the variable “adaptation process advancement” were calculated. In the last step of the research the language versions both included in Hofstede’s research, and having automatic welcoming messages were distinguished. The correlations between the elements of the messages offering encouragement, requirements, help, the complexity of the messages and the chosen Hofstede’s indices: femininity, uncertainty avoidance and individuality were calculated.

4. Results

To investigate the complexity and advancement of adaptation process designed for newcomers in all the language versions of Wikipedia, the tools used for adaptation have been counted. The results are presented in Table 1.

Table 1: Number and percentage of Wikipedia language versions using particular adaptation tools.

<table>
<thead>
<tr>
<th>Adaptation tool</th>
<th>Number of Wikipedias</th>
<th>% from 325 Wikipedias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Help page</td>
<td>273</td>
<td>84.8%</td>
</tr>
<tr>
<td>Village Pump</td>
<td>244</td>
<td>75.1%</td>
</tr>
<tr>
<td>FAQ page</td>
<td>114</td>
<td>35.1%</td>
</tr>
<tr>
<td>Introduction page</td>
<td>110</td>
<td>33.8%</td>
</tr>
<tr>
<td>Automatic Welcome Template</td>
<td>42</td>
<td>12.9%</td>
</tr>
<tr>
<td>Welcome templates category</td>
<td>34</td>
<td>10.5%</td>
</tr>
<tr>
<td>Teahouse</td>
<td>26</td>
<td>8.0%</td>
</tr>
<tr>
<td>Welcoming Committee</td>
<td>21</td>
<td>6.5%</td>
</tr>
</tbody>
</table>

As indicated in the collected data, two of the adaptation tools exist in the majority of Wikipedia language versions – the Help page (84.8%) and the Village Pump (75.1%). Around one-third of wikis have got frequently asked question pages (FAQ). Likewise, one-third of wiki versions created the Introduction page. What is more, only a small percentage of wikis have prepared automatic welcome message template (12.9%), welcome templates category (10.5%), Teahouse (8%) and a welcoming committee (6.5%).

The welcome templates used by the more experienced users to welcome and get acquainted with newcomers have been counted and analyzed in terms of differences. The exemplary variances of the templates are presented in Fig. 1.
Figure 1: Examples of welcome templates in Indonesian and English Wikipedia

The templates are not the same in all the language versions; they differ in terms of complexity, graphic elements used, a number of links (leading to the aspects of Wikipedians’ activity in the project) used. Some are really complex, visually attractive, while encouraging newcomers to edit, whereas some are more simplistic, minimalistic, including only the bare text.

With the aim of verifying H1 and H2, the author examined the correlations between the variable adaptation process advancement (AdaptToolSum) and the chosen variables characterizing Wikipedia language versions referring to the number of active users (ActiveUsersNo), the size of language versions output (ArticlesNo, TotalPagesNo, EditsNo, AdminsNo, RegisteredUsersNo), and the productivity of active users (ArticlesNo./ActiveUsersNo, TotalPagesNo./ActiveUsersNo, EditsNo./ActiveUsersNo). It was assumed that in versions with advanced and complex adaptation processes, the number of active users as well as the productivity of active users would be higher.

Table 2: Correlations between chosen variables characterizing Wikipedia language versions and the variable “adaptation process advancement”. N=325, p<0.05.

<table>
<thead>
<tr>
<th>Variable</th>
<th>ArticlesNo.</th>
<th>Total PagesNo.</th>
<th>EditsNo.</th>
<th>AdminsNo.</th>
<th>RegisteredUsersNo.</th>
<th>Active UsersNo.</th>
<th>Articles No./Active UsersNo.</th>
<th>Total Pages No./Active UsersNo.</th>
<th>Edits No./Active UsersNo.</th>
</tr>
</thead>
<tbody>
<tr>
<td>AdaptTool Sum</td>
<td>0.68</td>
<td>0.71</td>
<td>0.69</td>
<td>0.67</td>
<td>0.68</td>
<td>0.72</td>
<td>0.40</td>
<td>0.51</td>
<td>0.49</td>
</tr>
</tbody>
</table>

There is a statistically significant correlation between the variables measured with Spearman’s rho (p<0.5). The results illustrated in Table 2 reveal that in cultures with a higher level of adaptation process advancement, both the number of active users and their productivity are higher. The hypotheses 1 and 2 have been confirmed.
The scatter plot of data presenting the relationship between the sum of adaptation tools in a given language version and active users number is presented in Fig. 2. It presents how the active users number rises with the rise of the advancement of the adaptation process.

While counting automatic welcome templates, their similarity and complexity were analyzed. The conclusions were the same as in the case of the welcome templates used by experienced users to welcome newcomers (Fig. 1) – they differed in terms of the number of elements (links) included, information provided, visual attractiveness, and graphic elements used. The exemplary variances of the templates are presented in Fig. 4.

Figure 2: The relation between the variable AdaptToolSum and ActiveUsersNo. N=325.

To determine if H3, H3a, H3b, and H3c are true, the author calculated correlations between the chosen Hofstede indices – Individualism (IndivIndex), Femininity (MASIndex) and Uncertainty avoidance (UncertainIndex) and other chosen variables: the level of complexity of automatic welcome message template (measured with the sum of the elements presented – AutomaticWelcomeElemSum), the sum of elements in automatic messages encouraging the newcomers (AutomaticWelcomeEncourag), providing requirements (AutomaticWelcomeRules&Requir), and offering help (AutomaticWelcomeHelp). The calculated results are shown in Table 3.
Table 3: The correlations between the level of complexity of automatic welcoming message, encouragement, requirements and help in the automatic welcoming messages templates and chosen Hofstede indices value. N=15, p<0.05

<table>
<thead>
<tr>
<th>Variables</th>
<th>Spearman’s rank order correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Automatic Welcome ElemSum</td>
</tr>
<tr>
<td>IndivIndex</td>
<td>0.427496</td>
</tr>
<tr>
<td>MASIndex</td>
<td>-0.513542</td>
</tr>
<tr>
<td>UncertainIndex</td>
<td>-0.103209</td>
</tr>
</tbody>
</table>

We may note interesting findings that emerge from the data analysis. Three statistically significant correlations have been revealed. The masculinity index correlates negatively with the complexity of the automatic welcome message template (-0.51), as well as with the encouragement level (-0.67). The individuality index correlates positively with the level of rules and requirements of the automatic welcome message template (0.53). Unexpectedly, the encouragement and offering of help did not correlate significantly with the individualism-collectivism dimension. Only one aspect of the automatic welcome template content correlated with it, albeit the direction of the correlation was contrary to the expectations. As demonstrated by the results, some aspects of Wikipedia content are culturally dependent, but not all and not with every dimension of culture. The hypotheses 3a and 3c have been partly confirmed, while H3b has been rejected.

5. Discussion

Of all the main tools dedicated to the adaptation of newcomers in Wikipedia, only two exist in the majority of its language versions, namely the Help page (84.8%) and the Village Pump (75.1%), while one-third of wikis have got the frequently asked questions page (FAQ) and the Introduction page. Yet, some language versions of FAQ (e.g. Polish) have not been updated after successive changes of user interface (facilitating other ways of communication, other ways of editing pages, new kinds of user rights). This fact may reduce their usefulness to the newcomers. It is worth noting that a very small percentage of wikis use automatic welcome message template (12.9%), welcome templates category (10.5%), Teahouse (forum for newcomers) (8%), or a welcoming committee (6.5%). These tools are underestimated in the majority of Wikipedia language versions. It was proven by Morgan and Halfaker (2018) that retention of newcomers invited to the Wikipedia’s “Teahouse” was higher. Hence, it would be advisable to use this tool more widely or invite newcomers to the special section dedicated to them in public forum.

It was found by numerous researchers that the strategies representing institutionalized socialization were negatively related to intentions to quit, while also positively relating to job satisfaction, organizational commitment, and activity performance. Newcomers who experienced institutionalized socialization usually have more certainty about their new role, while feeling more accepted, committed. They are less likely to drop out, and their activity is more effective (Saks, Uggerslev, & Fassina 2007; Choi, 2010, p. 108). The tools described in the presented paper were found to be more individualised than institutionalised. They are individual, random, informal, and variable. However, mentoring may be perceived as institutional socialization, because it is serial (experienced users serve as role models), not disjunctive (no role models). Similarly, a welcoming committee as a permanent body of members dedicated to initiating communication with newcomers, is rather formal, which is more typical for institutional socialization. Development of the tools typical for the institutional socialization in Wikipedia would be recommended.

As Chao et al. (1994) point out, new members who acquire greater knowledge about the internal group policy, the jargon, or group goals and values usually have a higher level of involvement and satisfaction from participating in the project. Not exactly the same, but similar conclusions arise from the presented paper. H1 and H2 were confirmed in the research – in Wikipedia versions with advanced and complex adaptation processes, the number of active users and the productivity of active users would be higher. The number of active users and content amount per active user rises with the growing advancement of the adaptation process.

It was expected that all three Hofstede indices correlate with several aspects of the automatic welcome message template content (H3). Surprisingly, some of the indicators studied appeared to have the opposite direction of correlation than was assumed (the individualism index: 0.535421). The scope of providing requirements is positively correlated with the individuality level in a given culture. After more profound analysis, one can explain
this phenomenon by the fact that more individualistic countries provide the necessary guidelines for the activity mainly in the template, whereas in more collectivist countries the guidelines might be transferred more directly, during interpersonal interactions in forums and discussion pages. Kukowska and Skolik (2021, p. 465) also note that collectivistic cultures are usually high-context ones, therefore not all the rules have to be explicitly indicated and thoroughly explained.

Moreover, it was proven that the more masculine a country is, the narrower the scope of complexity of the welcome template and the less encouragement for participation among newcomers (-0.513542 and -0.671635 respectively). This has confirmed the assumption from H3a, however only partly, as it was not applied to the aspects of the scope of the requirements, and help. This stands at least partly in accordance with the Hofstede’s statements (2007) about the countries with higher femininity index, which are characterized by facilitating cooperation, relationships, and satisfying others’ needs.

6. Conclusions and limitations

It is essential to create proper collaborative environment for newcomers, and strategies and tactics supporting it. Such adaption tactics as sharing knowledge with newcomers, regular offers of assistance, welcome messages, positive feedback on a newcomer’s edits, and other ones, especially of an institutionalized nature, are all recommended to be implemented in online projects such as Wikipedia.

The wide range of possible adaptation tools is not used by the majority of language versions, as only two are applied in over 75% of versions. Initiating communication with new users in the form of automatic welcome template or welcome templates dedicated to the users are appreciated in the minority of versions. There is still a lot that could be improved in this respect. While developing strategies and tactics concerning Wikipedia’s adaptation of newcomers, it is significant to take into consideration the cultural differences and different cultural expectations of the users.

It would be recommended to expand the research and confirm some of the results on a wider research population in various cultural contexts. The results may be of use to different virtual communities of practice and organizations which face the problem of adapting newcomers in the online knowledge sharing, or other open-source projects, as well as supporting virtual team management.

References


