Is Interaction More Important Than Individual Performance? A Study of Motifs in Wikia

Thomas Arnold
Research Training Group AIPHES
Department of Computer Science
Technische Universität Darmstadt
www.aiphes.tu-darmstadt.de

Johannes Daxenberger
Ubiquitous Knowledge Processing Lab
Department of Computer Science
Technische Universität Darmstadt
www.ukp.tu-darmstadt.de

Iryna Gurevych
Ubiquitous Knowledge Processing Lab
Department of Computer Science
Technische Universität Darmstadt
German Institute for Educational Research and Educational Information
www.ukp.tu-darmstadt.de

Karsten Weihe
Algorithmic Group
Department of Computer Science
Technische Universität Darmstadt
www.algo.informatik.tu-darmstadt.de

ABSTRACT

Recent research has discovered the importance of informal roles in peer online collaboration. These roles reflect prototypical activity patterns of contributors such as different editing activities in writing communities. While previous work has analyzed the dynamics of contributors within single communities, so far, the relationship between individuals' roles and interaction among contributors remains unclear. This is a severe drawback given that collaboration is one of the driving forces in online communities. In this study, we use a network-based approach to combine information about individuals' roles and their interaction over time. We measure the impact of recurring subgraphs in co-author networks, so called motifs, on the overall quality of the resulting collaborative product. Doing so allows us to measure the effect of collaboration over mere isolated contributions by individuals. Our findings indicate that indeed there are consistent positive implications of certain patterns that cannot be detected when looking at contributions in isolation, e.g. we found shared positive effects of contributors that specialize on content quality over of quantity. The empirical results presented in this work are based on a study of several online writing communities, namely wikis from Wikia and Wikipedia.

Keywords

Wikia; Online collaboration; Online Communities; Informal Roles; Co-Author Networks; Motifs

WWW'17 Companion, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3053362

1. INTRODUCTION

Much has been said about the importance of collaboration and interaction in online communities and social networks [28, 21, 31]. In particular, online writing communities have attracted research in this regard due to their importance as public knowledge resources. Some studies claim that few experts involved in the collaborative process are crucial for a positive outcome [24, 23]. Others found evidence that many potentially small contributions by layman are the most important factor [30, 18]. Yet another stream of literature claims that coordination is crucial to leverage the wisdom of the crowd [5, 19]. In particular, recent research has highlighted the role of implicit coordination which emerges organically [16, 3]. In online writing communities, this kind of implicit coordination has been modeled in the form of informal roles, reflecting the editing history of contributors based on different kinds of edit actions they performed [3, 21]. For example, a contributor who frequently corrected typos and grammatical mistakes could be characterized with the informal role "copy-editor".

In isolation, informal roles reveal much about *what* contributors do, but little about *whom* they interact with while working. Given the high number of studies showing the importance of interaction in online collaboration [27, 10, 20], it is highly desirable to combine informal roles with detailed information about who interacts with whom. Our assumption is that the overall performance of an online community not only depends on the performance of single contributors or the number of contributors in total, but rather on the way they interact with each other – in particular, who interacts with whom. The intuition behind this assumption is that interaction between diverse types of contributors is more beneficial for the collaborative outcome. If, for example, copy-editors interact with contributors creating a lot of content, this could be favorable over the collaboration of content creators alone.

To test this assumption, we integrate a fine-granular analysis of edit activity and resulting implicit roles of contributors with a graph-based approach to measure interaction. We are thus able to not just quantify interaction in online communities, but we also describe the kind of interaction and the types of contributors interacting. To measure the influence of informal roles and contributor interaction on the knowledge production process, we further take the quality of the outcome into account. To be able to scale across various online communities, we chose to use the fan-based for-profit community Wikia as our main investigation target. Compared to Wikipedia[1] (launched in 2001, approx. 2 million active users), Wikia[2] is a rather restrictive community (launched in 2006, approx. 11,000 active users) with a clear commercial background and more editing limitations.

Our findings highlight substantial differences in the revision behavior of different online communities. While the open editing policy of Wikipedia results in a significant administrative overhead to prevent vandalism, it also helps to ensure sustainable collaborative structures and a balanced community of editors. We identify important interaction patterns ("motifs") which characterize but also distinguish the editing work across communities within Wikia. Our analysis suggests that a combination of contributors' informal roles and their interaction in terms of network motifs yields a consistent picture of community performance.

2. RELATED WORK

Previous work on collaboration in online and social networks has extensively analyzed the interaction between contributors based on graph structures [28, 31]. Most of these looked into quantitative properties of co-author networks or subgraphs. For example, Sachan et al. [25] analyze social graphs on Twitter and in email conversations to discover smaller communities of contributors with shared interests. Brandes et al. [10] define co-author networks to visualize differences in the behavior of contributors and to reveal polarizing articles in Wikipedia. Their networks are based on positive and negative interaction of Wikipedia contributors in the form of delete and undelete actions. These approaches have two limitations. First, they largely ignore the temporal dimension. A static analysis of graph structures, however, can only reveal limited insight, as online communities and particularly social networks tend to evolve dynamically [16, 15]. Second, as they are typically based on (social) links between contributors (such as followers, likes etc.), they do not take into account the informal roles of contributors. The latter might however reveal important information about the implicit coordination inside the network. Jurgens and Lu [15] address these concerns by integrating formal roles (e.g. admin, bot) and the temporal sequences of edits into their analysis of Wikipedia. With this approach they are able to identify four types of contributors' behavior with increasing or decreasing frequency over the course of time in Wikipedia's history. However, both their model of edit types as well as their model of contributor roles are pretty course-grained and capture rather high-level properties of the collaborative process.

Another stream of literature has analyzed informal (or social) roles in online communities. As opposed to formal roles [6], informal roles are not awarded by an authority, but they emerge organically. For example, the posting behavior of contributors on reddit has been used to identify roles such as the "answer-person" [11]. Welser et al. [29] describe four social roles played by Wikipedia contributors, based on a small-scale manual analysis of edit patterns and a larger-scale analysis of edit locations. They find that new contributors would quickly adapt to fit into one of those roles and that their notion of social roles implicitly models the "social" network of contributors, i.e. their interaction on Wikipedia talk pages. In our approach, we adopted a slightly different notion of informal roles, based on contributors' edit history. It involves a fine-granular classification of Wikipedia edit types such as spelling corrections, content deletion or insertion [12]. This method has first been suggested and tested for the online community Wikipedia by Liu et al. [21] and improved by subsequent work. Among the latter, Yang et al. [33, 34] present a method for automatic multi-label classification of Wikipedia edits into 24 types, based on a manually annotated sample. They identified eight roles based on editing behavior, involving a manual evaluation. The training data for edit categories used by Yang et al. [34] is rather small, and the performance of their automatic edit classification algorithm is lower as compared to the revision-based classification approach presented by Arazy et al. [3] and used in this work.

With respect to the analysis of co-author networks, our work builds upon Brandes et al. [10]. However, in contrast to Brandes et al., we use informal roles of contributors to create more generalized networks, which enables us to search for universal interaction motifs. The exploitation of network motifs for analyzing collaboration in Wikipedia has previously been proposed by Jurgens and Lu [15] and Wu et al. [32]. The latter used their analysis of motifs to predict article quality. In a similar vein, Arnold et al. [7] construct sentence-level networks based on shared nouns to predict high quality Wikipedia articles. In addition, they analyze the linguistic connection between the most dominant motifs and text quality. The approach proposed in this work is different from previous work on motif analysis in online collaboration in that we measure the impact of recurring motifs based on informal roles for entire communities rather than single articles.

3. ONLINE COLLABORATION IN WIKIA

Wikis offer a convenient resource to study collaborative writing behavior as they have low entry barriers for new contributors, but at the same time they offer a reasonable administrative structure which allows to record and reverse any editing action. In the present study, we analyze wikis from the wiki hosting service Wikia. Wikia is a hosting service for wikis with a focus on entertainment fan sites.[3] Its users are not charged for creating wikis, contributing or accessing information. Nevertheless, the operator Wikia Inc. is a for-profit company, and it generates profit from Wikia in the form of advertisement. In contrast to the broad scope of topics in the online encyclopedia Wikipedia, the main focus of Wikia is entertainment. Most Wikia communities cover topics from television, movie or (computer) game genres. Overall, Wikia hosts over 360,000 communities with over 190 million unique visitors per month.[4]

[1] `https://stats.wikimedia.org`

[2] `http://community.wikia.com/wiki/Special:ListUsers`

[3] In October 2016, Wikia.com has been renamed to "Fandom powered by Wikia" to strengthen the association with the "Fandom" brand.

[4] `http://www.wikia.com/about`

As opposed to Wikipedia, where the internal quality rating of articles follows a strict process, there are no global quality estimators for Wikia articles.[5] Since January 2012, Wikia provides a combined indicator of performance, traffic and growth for every individual community – the Wikia Activity Monitor (WAM).[6] This single score between 0 and 100 is recalculated on a daily basis, and is used to rank the communities. To prevent aimed manipulation of this score, the specific formula is not known to the public. As this score is applicable and comparable across Wikia communities, we use the WAM score as a global measure of community performance.

For our experiments, we chose a selection of seven English Wikia communities, based on high WAM score, reasonable size and genre diversity (see Table 1). More specifically, we excluded all Wikia communities that either (a) are non-english, (b) have too unusual structure, like lyrics or answers, (c) have over 200,000 revisions and would therefore require very long computation time, (d) did not have an available database dump from January 2016 or newer or (e) have a WAM score below 85. From the remaining choices, we select the five communities with highest revision count: Disney, Tardis (TV series), WoW (World of Warcraft – video game), Villains and The Walking Dead (TV series). Since all five have very high WAM scores over 97, we handpicked two additional communities: "24" as an additional TV Series wiki with a WAM score of 86, and Military, which has a WAM score of 85, for additional genre diversity.

4. METHODOLOGY AND RESULTS

The following section provides an overview of our approach and explains essential principles. We utilize automatic classification of revision categories (Section 4.1) and consequently determine informal roles for contributors in writing communities (Section 4.2). We then create a network based on individual contributor interaction and use a novel contributor role model to extract general collaboration patterns (Section 4.3). These patterns yield insight about similarities and differences of wiki communities, and we explore the effect of these patterns on the success of a community. To access and process data from Wikipedia and Wikia, we used the freely available Java Wikipedia Library [14]. Our results are based on the June 2015 database dump from the English Wikipedia and March 2016 dumps from Wikia. For the sake of readability, we present each method and result stepwise.

4.1 Revision Classification

We focus our research on contributor interaction in collaborative platforms based on writing processes. Contributors create online articles, and these articles are extended and refined by the same or other contributors. Every revision serves one or more purposes, such as adding content, spelling corrections or adding citations. Additionally, changes from one contributor can be completely revoked by another contributor. We classify revisions with the edit-based multi-label classification method proposed by Daxenberger et al. [13] that has later been adapted to revision-level by Arazy et al. [3]. As training data, we use the data set described by the same article [3] with more than 13,000 manually labeled Wikipedia revisions and twelve revision types, such as "Add Substantive New Content", "Rephrase Existing Text" or "Add Vandalism" (full list see Table 2). A detailed description of the revision types can be found in [3] and [2]. This training data is used in a machine learning setup to create a model for automatic prediction of revision types on unseen data. Following Arazy et al. [3], we use a set of manually crafted features based on grammatical information (e.g. the number of spelling errors introduced or deleted), meta data (e.g. whether the author of the revision is registered), character- and word-level information (e.g. the number of words introduced or deleted) and wiki markup (e.g. the number of internal links introduced or deleted) of each revision. This information is then used by a Random k-Labelsets classifier [26], an ensemble method which optimizes the output of several decision tree classifiers, to classify revisions. The proposed method yields state-of-the-art performance on the Wikipedia dataset from Arazy et al. [3].

We applied the classification to a large number of Wikipedia and Wikia revisions, as listed in Table 1. The performance of this classification of Wikipedia revisions has been shown previously [3], but we apply this method on Wikia revisions, using the same training data from Wikipedia. As we did not know about the effect of this change of domain (training on Wikipedia revisions, testing on Wikia revisions), we did a small-scale manual evaluation on Wikia data. Based on a manual evaluation of 100 random revisions from our Wikia sample, the classification of Wikia revision yields results comparable to Wikipedia revision classification with 0.66 macro-F1 as compared to 0.68 in Wikipedia as reported by Arazy et al. [3].

Results.

Table 2 shows the distribution of the twelve revision classes on the Wikipedia sample and the seven Wikia communities. The distribution allows first observations on similarities and differences of wiki communities. Compared to the Wikia communities, the Wikipedia data set has a higher share of "Add Vandalism" and "Delete Vandalism" revisions. Since Wikipedia attracts a much larger and broader audience as compared to Wikia, it also attracts more misbehavior, which results in the need of explicit counter-measures against these destructive actions.

Comparing the values of the seven Wikia communities shows their heterogeneous nature. The Wikia communities with a higher revision-to-page ratio, like Walking Dead, Tardis and WoW, are quite similar to each other and to the Wikipedia data. In contrast, Villains and Military, which both have a very low revision-to-page ratio, show significant differences. The share of "Reorganize Existing Text" revisions in Villains is more than twice as high as in every other data set. Military has an exceptionally high share of "Create a New Article" revisions, which is reasonable, given that it has by far the lowest amount of revisions per article. Furthermore, it seems to attract a high proportion of unusual edits, as shown by the above-average number of "Miscellaneous" revisions. Our findings indicate that maturity (as measured by the number of edits, as well as by age) influences revision behavior in online writing communities. Motivated by this finding, in the next section we go

[5] Although there are panels of contributors rating individual articles in some Wikia communities, there are no overarching norms for quality control across all Wikia wikis.

[6] http://www.wikia.com/WAM/FAQ

	Disney	WoW	24	Tardis	Villains	The Walking Dead	Military	Wikia (sample)	Wikipedia (sample)
Revisions	158,733	122,449	56,509	126,318	105,273	105,138	75,028	107,064	877,717
Pages	1,710	1,148	914	564	2,323	425	13,189	2,896	1,000
Ratio	92.82	106.66	61.826	223.96	45.31	247.38	5.68	111.94	877.72

Table 1: Basic statistics of our data sets.

	24	Disney	Military	Tardis	Villains	Walk. Dead	WoW	Wikia Average	Wikipedia
Add Citations	1.18	1.11	2.22	1.47	0.54	1.12	2.90	1.51	1.84
Add New Content	21.99	20.59	7.03	19.94	10.73	23.36	22.33	18.00	22.29
Add Wiki Markup	26.93	30.49	36.31	27.72	36.46	24.91	27.78	30.09	28.09
Create a New Article	0.52	0.25	6.32	0.18	0.72	0.06	0.32	1.20	0.42
Delete Content	11.16	8.60	3.25	9.49	4.16	10.63	10.64	8.28	6.64
Fix Typo(s)/Gramm. Err.	12.88	11.46	22.10	16.28	10.88	12.96	11.93	14.07	12.23
Reorganize Existing Text	9.82	13.21	12.43	10.14	28.42	7.44	9.25	12.96	4.59
Rephrase Existing Text	4.43	5.40	1.16	5.22	2.96	6.09	4.38	4.23	3.88
Add Vandalism	8.80	6.45	1.18	6.79	4.10	9.03	7.72	6.30	9.93
Delete Vandalism	1.35	2.10	2.81	1.51	0.77	3.94	1.80	2.04	7.85
Hyperlinks	0.17	0.12	0.59	0.34	0.06	0.18	0.15	0.23	1.50
Miscellaneous	0.78	0.22	4.61	0.94	0.20	0.29	0.80	1.12	0.75

Table 2: Revision type distribution of different wiki communities, in percent.

one step further and turn the revision behavior of individual contributors into a set of roles, which characterize the writing process in the entire community.

4.2 Informal Roles

In order to define generic motifs interaction (rather than individual editor collaborations), the individual contributors of both collaborative platforms – Wikipedia and Wikia – have to be mapped to a fixed set of roles that are based on revision types. Contributors with similar writing behavior in the context of a specific community should be assigned to the same informal role. We create revision type vectors for every contributor in every article, using the results of our revision type classification. Each vector contains the revision type frequency of every revision the contributor created for a given article, normalized to sum up to 1. We detect informal roles from all vectors through a k-means clustering algorithm, with the number of clusters k varying between 2 and 10.[7] We compare the results via Overall Cluster Quality (OCQ) values, which is a balanced combination value of cluster compactness and cluster separation [3, 21]. The clustering with best OCQ values was chosen as the best informal role representation. For the 1,000 Wikipedia articles sample, this results in seven roles as described in previous work by Arazy et al. [3].

For our Wikia data sets, we considered two different approaches to cluster the contributors. The first strategy involves individual clustering of every Wikia community. As for Wikipedia, we used the same k-means clustering approach. From these possible clusterings, we selected the best option based on OCQ values. With this method, we were not able to create comparable informal roles across multiple Wikia communities (nor comparable to the ones we found for Wikipedia), which made it very hard to detect general collaboration patterns. The results can still be useful to compare Wikia communities, but the clusters from different Wikias are too diverse to draw meaningful conclusions across community borders, which makes this first approach unfeasible.[8] Therefore, we decided to map all Wikia contributors to a single, shared set of common roles, based on one global clustering on a combined data set of all Wikia revisions. We expect that a meaningful global role mapping for many Wikia communities might require a different number of clusters than Wikipedia. We considered all possibilities between 2 and 15 clusters. From these options, we selected the final clustering for all Wikia communities based on optimal OCQ values.

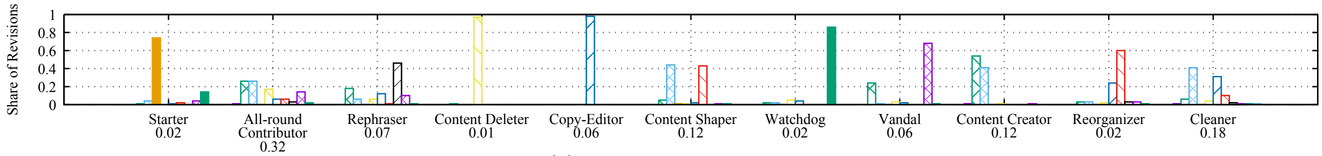
Results.

The best clustering for Wikia is presented in Figure 1a. It contains eleven informal roles: *Starter* (focus on creating new articles), *All-round Contributor* (no particular focus), *Rephraser* (focus on rephrasing content and adding text), *Content Deleter* (focus on deleting content), *Copy-Editor* (focus on fixing typos), *Content-Shaper* (focus on organizing content and markup), *Watchdog* (focus on vandalism detection), *Vandal* (focus on adding vandalism), *Content Creator* (focus on adding content and markup), *Reorganizer* (focus on moving text and fixing typos), and *Cleaner* (focus on fixing typos and markup). A comparison to the Wikipedia cluster centroids discovered by Arazy et al. [3] (cf. Figure 1b) reveals some similarities, but also characteristic differences. One key similarity can be observed in the biggest cluster of each respective data set. These clusters both have a strong "Allround"-character, as their class distribution vectors have no clearly dominant dimension. This indicates that the majority of contributors does not focus on one single type of task.

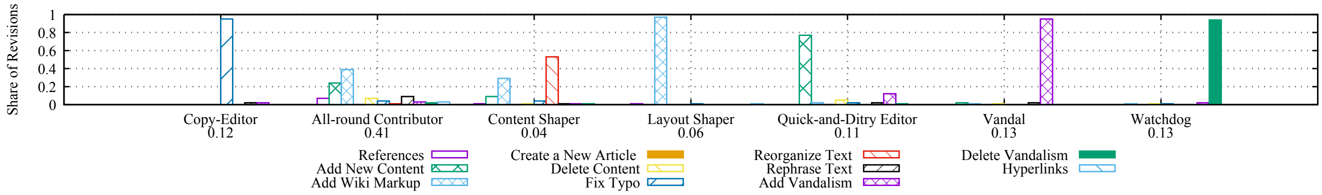
The Wikia clustering contains several distinctive roles. Among these is the "*Starter*" role with a very large share of "New Article" revisions. Many Wikia articles only attract few edits after their creation, so an informal role that is limited to the creation of new articles is more likely. Communities with comparatively low number of revisions (cf.

[7]Following Arazy et al. [3], we used the k-means++ method [8] as initialization for the clusters and tested a range of random seeds.

[8]Please note that this finding adds to previous work, which found that the nature of informal roles in Wikipedia remains stable over time [3]. Our results suggest that – while within communities stable informal roles do exist – this is not necessarily the case in different communities.



(a) Wikia informal roles.



(b) Wikipedia informal roles.

Figure 1: Global distributions of informal roles (fraction of contributors per cluster below role names) for our samples.

revision to page ratios in Table 1) – like "Military" – always contain an informal role with a strong focus on "New Article" revisions. The "*Content Deleter*" role is also unique to the context of Wikia, and contains contributors almost exclusively shortening and deleting content. Furthermore, we detected a role with contributors focusing on both adding markup and fixing typos. Its scope is a bit broader as compared to the "*Spelling Corrector*" role in Wikipedia.

4.3 Collaboration Patterns

Having identified the different roles played by Wikipedians and Wikia contributors, we can analyze the interactions between types of contributors. Therefore, we use article-based co-author networks, in which contributors form nodes and interactions between contributors form directed edges [10]. We calculate such a network for each article from our Wikia sample. We map all contributors to the informal role they played in a particular article. We then count all interactions, across all articles, between the different contributors. Lastly, we analyze the effect of general interaction sub-patterns ("motifs") on community performance (in Wikia). In the following, we describe this process in more detail.

4.3.1 Co-author Networks

Brandes et al. [10] propose an edit network based on sentence-level interaction. In their network, each Wikipedia article forms a graph $G = (V, E)$. The nodes V correspond to the contributors who have performed at least one revision. The directed edges $E \subseteq V \times V$ represent interaction between a pair of contributors. As our intention is to understand collaboration between contributors based on their informal roles, we decided to slightly simplify the original co-author network of Brandes et al. [10]. In Contrast to Brandes et al. [10], we define the following types of interaction for a pair of contributors $u, v \in V \times V$: a) u **supports** v, and b) u **deletes** v. The support interaction indicates that contributor u changed or added information to a sentence that contributor v has created or edited. If contributor u completely removes a sentence written by contributor v or reverts that contribution, we create a delete interaction. After the full network is created, we replace the labels of all nodes V – the individual contributor – with their respective informal role, according to Section 4.2. See Figure 2

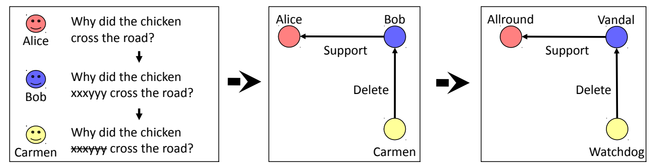


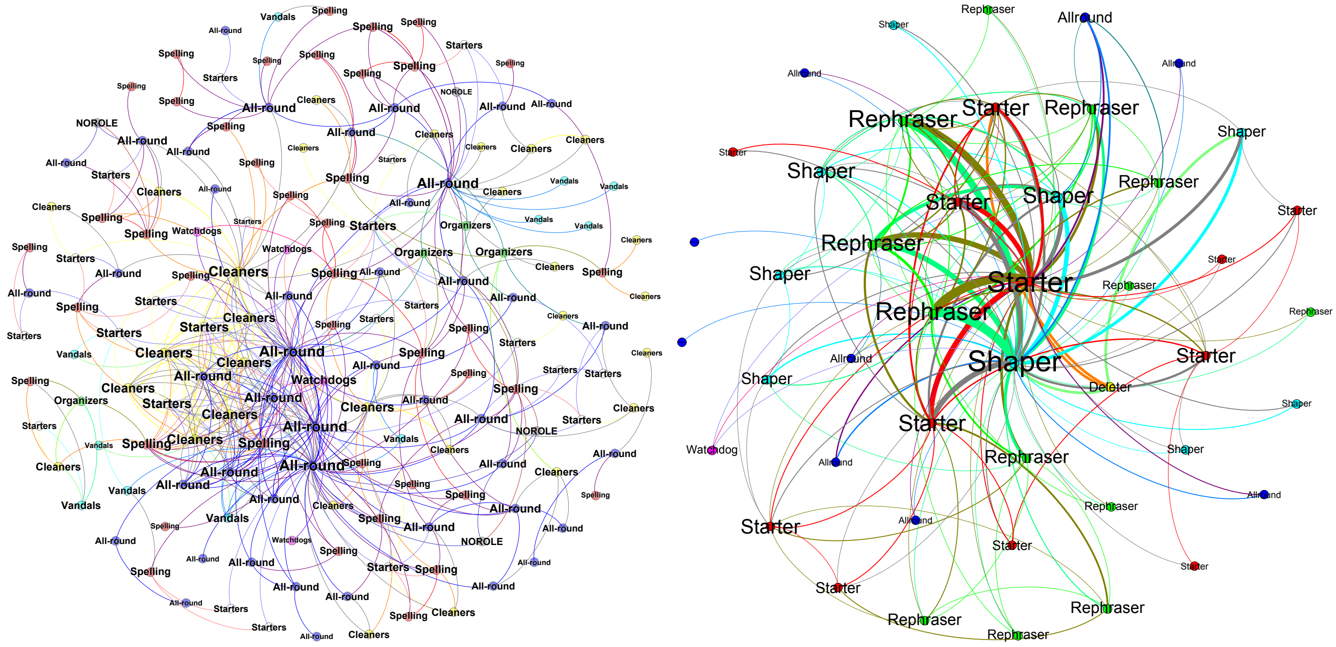
Figure 2: Example for co-author network creation. First step: Identify sentence-level edits. Second step: Create network with support and delete interactions. Third step: Replace contributor identification with respective informal role.

for a small example, and Figure 3 for visualizations of two support networks.

4.3.2 Motifs

Based on the simplified co-author network, we identify recurring collaboration patterns in the network – so called motifs [22]. These are defined as repeated interactions of the same type within the same edit context. As the delete interactions cannot be repeated within the same context – contributor v adds or edits content, contributor u reverts it – delete interactions are already interaction chains of maximum length. In contrast, support interactions can form chains of any length. If, for example, contributor v adds content and contributor u edits some of it and adds more information, the resulting interaction chain would be: u supports v. Then, a third contributor w adds some wiki markup in the same context, the resulting interaction chain would be: w supports u and v. To identify the basic motifs of the rather long interaction chains, they are split into pairs. As the example in Figure 4 indicates, the interaction chain "*All-round Contributor* supports *Starter* and *Copy-Editor*" is split into the two motifs "*All-round Contributor* supports *Starter*" and "*All-round Contributor* supports *Copy-Editor*". We consider these interactions motifs to be the building blocks of collaborative interaction.

We identify motifs of noticeable high frequency by comparison to randomly generated null-models [15], based on the interaction chains of informal roles. To generate a null-model, we keep the length and frequency of all interaction chains, but remove its informal role labels. This gives us the distribution of informal roles and a basic structure of interactions, with support chains of different length and frequency and a number of delete interaction pairs. We then



(a) (b)

Figure 3: Graph visualization of Support interactions in the Wikipedia article 'Abscess' (a) and the Disney Wikia article 'R2D2' (b). The nodes are single contributors, identified by their informal role, and the edges indicate interaction between two contributors. The size of the nodes and edges reflect the number of contributions and frequency of interaction.

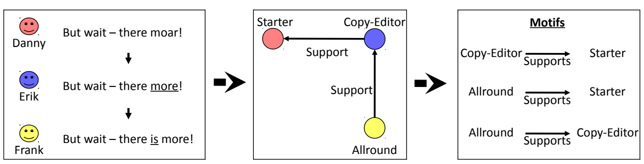


Figure 4: Example for interaction chains and pairwise motifs. The *Copy-Editor* supports the *Starter*. The *Allround contributor* further expands the combined work of both previous editors. This results in two additional support motifs.

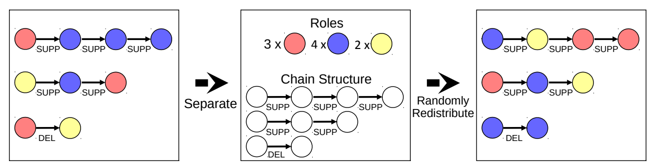


Figure 5: Example for null-model creation. We remove the role labels from all interaction chains. Then, we redistribute the labels randomly for each null-model.

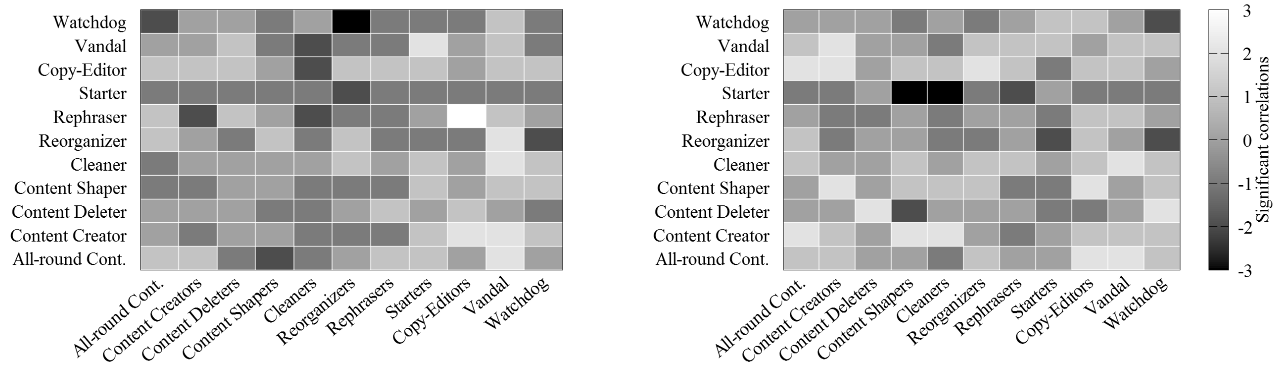
redistribute the informal roles randomly to this structure. In this manner, we get the exact same chain lengths, same distribution of support and delete actions, and the same distribution of roles, but potentially different motifs. Figure 5 displays an example of the null-model creation process with two support chains, one delete chain and three different informal roles.

We create 1000 random null-models for every collaborative community. Based on these, we calculate the z-score of every support and delete motif as $z = \frac{F_G(G') - \mu_R(G')}{\sigma_R(G')}$, where $F_G(G')$ is the frequency of a given motif in our data. $\mu_R(G')$ and $\sigma_R(G')$ indicate mean frequency and standard deviation of that motif in the randomly generated null-models. The z-score compares one value of a group of values to the mean [17]. In our case, high z-score values imply a remarkably high count of a particular motif compared to random chance. For example, the motif "*Rephraser* supports *Copy-Editor*" is found 5566 times in our Disney Wikia snapshot of January 2016. In our 1000 random null-models, the mean frequency of this motif is 4783.08 with a standard deviation of 43.69, which results in a z-score of 16.02. Since this is a relatively large positive number, it indicates that this motif is much more frequent in our real data in comparison to the random models.

Finally, we want to analyze the effect of unusual high or low frequency of specific motifs on the overall performance of the community. We conduct this analysis for every Wikia community, and use the WAM-score as an indicator of community performance. Since Wikia started publishing WAM-scores in January 2012, we consider seven points in time for our experiments in a 6-month rhythm, from January 2012 to January 2016. We determine the correlation between motif z-scores and the respective WAM score at each point in time with the Pearson correlation coefficient. In our case, a correlation coefficient of 1 for a specific motif would mean that a linear increase in the WAM score corresponds to a linear increase of the z-score of the motif. A correlation coefficient of -1 indicates linear negative correlation, where a linear increase in the WAM score corresponds to a linear decrease of the motif's z-score. If there is no correlation between z-score and WAM score, the correlation coefficient is 0.

Results.

Our motif research is based on interaction graphs that contain all support or delete interactions between contribu-



(a) Support motifs.

(b) Delete motifs.

Figure 6: Heatmap of correlation between motifs and Wikia WAM score. Light / dark color indicates a number of Wikia communities that showed statistically significant positive / negative correlations of the motif and Wikia WAM Score.

tors in a single article. Figure 3 features visualizations of two prototypical graphs from one Wikipedia and one Wikia article. As seen in the graphs, the collaborative writing process in Wikia (Figure 3b) is much more centralized, as most of the interaction involves a small group of main contributors. These central persons can also be seen in the Wikipedia article graph (Figure 3a), but there are also small teams and subgroups that do not necessarily involve the main contributors. This difference is indicated by the Louvain modularity measure [9]. In the given example, the modularity of the Wikipedia graph is five times as high as the modularity of the Wikia graph (0.519 versus 0.101).

To identify and illustrate the most important motifs in our Wikia communities, we combine the significant positive and negative correlations across all Wikia communities into a single heat map. Figure 6 depicts the results for both support and delete interaction motifs. Light color (up to white) indicates positive correlation of the respective motif and the Wikia WAM score, dark color (up to black) indicates negative correlation. The support heat map shows a strong positive effect of the "*Rephraser* supports *Copy-Editor*" motif. Support interactions of similar roles are also positively correlated with the Wikia WAM score, like "*Reorganizer* supports *Content Shaper*" or "*Copy-Editor* supports *Reorganizer*". All these roles focus on small corrections and quality improvements, rather than the creation of new content. In contrast, the "*Content Creator* supports *Content Creator*" interactions shows slightly negative effects on the success of the community. The *Content Creators* role includes contributors that mostly focus on adding more content. This is an additional indication for the importance of quality improvements over quantity improvements.

The support motif "*Watchdog* supports *Reorganizer*" has the highest negative correlation with the Wikia WAM score. Almost all support interactions of the "*Watchdog*" role have negative values in the heat map. In contrast, the delete motifs heatmap show that delete interactions of "*Watchdog*" have more positive effects, which confirms that the main focus of this informal group should be on removing potentially problematic content. Delete interactions targeting the "*Vandal*" role are strongly correlated to high community success. All support and delete motifs from the "*Starter*" role have negative correlation coefficients.

	Corr.	SD
Delete Substantive Content	-0.40	0.57
Fix Typo(s)/Gramm. Errors	0.39	0.52
Rephraser	-0.49	0.46
Cleaner	0.42	0.59
Cleaner supports Vandal	0.68	0.17
Allround Contr. supports Content Creator	0.59	0.28

Table 3: Mean correlation coefficient and standard deviation of the revision types, informal roles, and motifs with highest absolute correlation to Wikia WAM score.

5. DISCUSSION

Arazy et al. [3] showed that the nature of informal roles, i.e. the result of clustering contributors, in Wikipedia did not differ much across two periods of time. They conclude that the set of informal roles they discovered shows a high stability within the online community Wikipedia. However, when comparing communities in Wikia, we found that the nature and maturity of a writing community might well have an influence on informal roles, and consequently, contributor interaction. The differences could be the result of the fact that Wikia is less restrictive with regard to its content. For example, Wikipedia follows the principle of the "Five Pillars"[9], whereas wikis on Wikia are not bound to an encyclopedic content and format. Wikipedia's principles offer a "boundary infrastructure" [3] which Wikia lacks. A lack of such collaborative principles results in more nuanced and less stable collaborative structures, indicated by the significant differences between individual informal role clusterings in Wikia and Wikipedia. As exemplified by the graphs in Figure 3, Wikia articles tend to evolve around a central incubator, interacting with contributors working on the quality rather than the content of the article. To the opposite, in Wikipedia, the collaborative process develops around a set of central contributors, who are dealing with all aspects of editing work.

Our main findings connect motifs to community performance of Wikia platforms. The heatmaps in Figure 6 indicate that certain interactions between contributors have a

[9] `https://en.wikipedia.org/wiki/Wikipedia:Five_pillars`

significant impact on the overall performance of the community. To verify that the interaction of contributors is indeed key to success (or failure), we also tested the correlation of occurrences of revision types and informal roles with the Wikia WAM score, using the same dataset as in our motif experiments. As indicated in Table 3, the revision types "*Delete Substantive Content*" and "*Fix Typo(s) / Grammatical Errors*" have the greatest effect on the WAM score of all Wikia communities. As for informal roles, "*Rephraser*" and "*Cleaner*" show high positive and negative correlation coefficients. However, looking at the most significant interaction motifs, we see that they both show higher mean correlation coefficients and lower standard deviations across the different Wikia platforms. This shows that interaction is a more reliable predictor of community performance as compared to mere editing behavior or informal roles.

Looking at the motifs in more detail, we did find generally more positive influence of roles focusing on smaller quality improvements such as formatting and fixing typos. This is most noticeable in the roles "*Copy-Editor*" and "*Cleaner*" that are often positively correlated with the Wikia WAM scores. In other words, successful Wikia communities tend to place more value on content quality instead of quantity. In contrast, interaction of informal roles that are more concerned with adding or removing content, like "*Starter*" and "*Content Deleter*", did show negative or neutral effects on the community. Our finding adds to the work of Daxenberger and Gurevych [12], who found that high quality articles in Wikipedia attract more "surface" edits rather than revisions dealing with content extension or modification. Thus, our study confirms that their finding is valid for collaborative online communities other than Wikipedia. As a consequence, wiki organizers and administrators should emphasize the importance of both diversity and interaction among contributors, and incorporate this in their internal structures and processes. A potential application which would benefit from this analysis is e.g. online team formation [1], where contributors with different information roles need to be brought together in the right way.

6. CONCLUSION

In this work, we combined measures of implicit coordination with those from contributor interaction to assess community performance. To this end, we analyzed contributors' informal roles on two popular wiki platforms, namely Wikipedia and Wikia. While informal roles help to estimate what contributors do, patterns of interaction from co-author networks reveal who they are working with. Rather than using collaboration patterns to detects trends [15], we leverage informal roles to analyze the effect of interaction on community performance. This approach helped to identify collaboration patterns with consistent positive or negative effect, which is not possible when looking at editing behavior or informal roles in isolation. Our results reveal a particularly positive influence of contributors with a focus on small contributions for text quality improvement. This finding, in combination with the more diverse collaboration patterns we found in different Wikia wikis, points to a clear need for measures to increase implicit coordination and quality assurance in public wikis by bringing together the right people [1].

We see several directions for future work. First, it might be very helpful to get insights about contributors' motivation. Recent work [4] revealed that changes in the implicit coordination of contributors can be linked to different motivational orientations. This dimension is absent from our current study. Another limitation of our approach is that we had to rely on the somewhat obscure WAM score as an indicator for community performance. Future work might look into more transparent measures, e.g. by assessing the quality of all articles in a wiki.

Acknowledgments

This work has been supported by the German Research Foundation as part of the Research Training Group "Adaptive Preparation of Information from Heterogeneous Sources" (AIPHES) under grant No. GRK 1994/1 and by the German Federal Ministry of Education and Research (BMBF) under the promotional reference 01UG1416B (CEDIFOR). The authors would like to thank Saurabh Shekhar Verma for conducting experiments in preparation for this work.

7. REFERENCES

[1] A. Anagnostopoulos, L. Becchetti, C. Castillo, A. Gionis, and S. Leonardi. Online Team Formation in Social Networks. In *Proceedings of the 21st International Conference on World Wide Web*, pages 839–848, Lyon, France, 2012.

[2] J. Antin, C. Cheshire, and O. Nov. Technology-Mediated Contributions: Editing Behaviors Among New Wikipedians. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, pages 373–382, Seattle, WA, USA, 2012.

[3] O. Arazy, J. Daxenberger, H. Lifshitz-Assaf, O. Nov, and I. Gurevych. Turbulent Stability of Emergent Roles: The Dualistic Nature of Self-Organizing Knowledge Co-Production. *Information Systems Research. Articles in Advance*, 2016.

[4] O. Arazy, H. Lifshitz-Assaf, O. Nov, J. Daxenberger, M. Balestra, and C. Cheshire. On the "How" and "Why" of Emergent Role Behaviors in Wikipedia. In *Proceedings of the 20th Conference on Computer-Supported Cooperative Work and Social Computing*, page to appear, Portland, OR, USA, 2017.

[5] O. Arazy, O. Nov, and N. Oded. Determinants of Wikipedia Quality: the Roles of Global and Local Contribution Inequality. In *Proceedings of the ACM 2010 Conference on Computer Supported Cooperative Work*, pages 233–236, Savannah, GA, USA, 2010.

[6] O. Arazy, F. Ortega, O. Nov, L. Yeo, and A. Balila. Functional roles and career paths in wikipedia. In *Proceedings of the 18th Conference on Computer Supported Cooperative Work and Social Computing*, pages 1092–1105, Vancouver, BC, Canada, 2015.

[7] T. Arnold and K. Weihe. Network motifs may improve quality assessment of text documents. In *Proceedings of TextGraphs-10: the Workshop on Graph-based Methods for Natural Language Processing*, pages 20–28, San Diego, CA, USA, 2016.

[8] D. Arthur and S. Vassilvitskii. K-means++: The Advantages of Careful Seeding. In *Proceedings of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1027–1035, New Orleans, LA, USA, 2007.

[9] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10):P10008, 2008.

[10] U. Brandes, P. Kenis, J. Lerner, and D. van Raaij. Network Analysis of Collaboration Structure in Wikipedia. In *Proceedings of the 18th International Conference on World Wide Web*, pages 731–740, Madrid, Spain, 2009.

[11] C. Buntain and J. Golbeck. Identifying Social Roles in Reddit Using Network Structure. In *Proceedings of the 23rd International Conference on World Wide Web*, pages 615–620, Seoul, Republic of Korea, 2014.

[12] J. Daxenberger and I. Gurevych. A Corpus-Based Study of Edit Categories in Featured and Non-Featured Wikipedia Articles. In *Proceedings of the 24th International Conference on Computational Linguistics*, pages 711–726, Mumbai, India, 2012.

[13] J. Daxenberger and I. Gurevych. Automatically Classifying Edit Categories in Wikipedia Revisions. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 578–589, Seattle, WA, USA, 2013.

[14] O. Ferschke, T. Zesch, and I. Gurevych. Wikipedia revision toolkit: Efficiently accessing wikipedia's edit history. In *Proceedings of the ACL-HLT 2011 System Demonstrations*, pages 97–102, Portland, OR, USA, 2011.

[15] D. Jurgens and T.-c. Lu. Temporal Motifs Reveal the Dynamics of Editor Interactions in Wikipedia. In *Proceedings of the 6th International Conference on Weblogs and Social Media*, pages 162–169, Dublin, Ireland, 2012.

[16] G. Kane, J. Johnson, and A. Majchrzak. Emergent Life Cycle: The Tension Between Knowledge Change and Knowledge Retention in Open Online Coproduction Communities. *Management Science*, 60(12):3026 – 3048, 2014.

[17] B. R. Kirkwood and J. A. C. Sterne. *Essentials of Medical Statistics.* Blackwell Scientific Publications, 1988.

[18] A. Kittur, E. H. Chi, B. Pendleton, B. Suh, and T. Mytkowicz. Power of the Few vs. Wisdom of the Crowd: Wikipedia and the rise of the bourgeoisie. In *Proceedings of the 25th Annual ACM Conference on Human Factors in Computing Systems*, San Jose, CA, USA, 2007.

[19] A. Kittur and R. E. Kraut. Harnessing the Wisdom of Crowds in Wikipedia: Quality through Coordination. In *Proceedings of the ACM 2008 Conference on Computer Supported Cooperative Work*, pages 37–46, San Diego, CA, USA, 2008.

[20] D. Laniado and R. Tasso. Co-authorship 2.0: Patterns of Collaboration in Wikipedia. In *Proceedings of the 22nd ACM Conference on Hypertext and Hypermedia*, pages 201–210, Eindhoven, The Netherlands, 2011.

[21] J. Liu and S. Ram. Who does what: Collaboration patterns in the wikipedia and their impact on article quality. *ACM Transactions on Management Information Systems*, 2(2):11:1–11:23, 2011.

[22] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. Network Motifs: Simple Building Blocks of Complex Networks. *Science*, 298(5594):824–827, 2002.

[23] J. Onrubia and A. Engel. Strategies for Collaborative Writing and Phases of Knowledge Construction in CSCL Environments. *Computers & Education*, 53(4):1256–1265, dec 2009.

[24] R. Priedhorsky, J. Chen, S. K. Lam, K. Panciera, L. Terveen, and J. Riedl. Creating, Destroying, and Restoring Value in Wikipedia. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work*, pages 259–268, Sanibel Island, FL, USA, 2007.

[25] M. Sachan, D. Contractor, T. A. Faruquie, and L. V. Subramaniam. Using Content and Interactions for Discovering Communities in Social Networks. In *Proceedings of the 21st International Conference on World Wide Web*, pages 331–340, Lyon, France, 2012.

[26] G. Tsoumakas, I. Katakis, and I. Vlahavas. Random k-Labelsets for Multi-Label Classification. *IEEE Trans. on Knowledge and Data Engineering*, 23(7):1079–1089, 2011.

[27] F. B. Viégas, M. Wattenberg, and K. Dave. Studying Cooperation and Conflict Between Authors with History Flow Visualizations. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 575–582, Vienna, Austria, 2004.

[28] G. Wang, K. Gill, M. Mohanlal, H. Zheng, and B. Y. Zhao. Wisdom in the Social Crowd: An Analysis of Quora. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 1341–1352, Rio de Janeiro, Brazil, 2013.

[29] H. T. Welser, D. Cosley, G. Kossinets, A. Lin, F. Dokshin, G. Gay, and M. Smith. Finding Social Roles in Wikipedia. In *Proceedings of the 2011 iConference*, pages 122–129, Seattle, WA, USA, 2011.

[30] D. M. Wilkinson and B. A. Huberman. Cooperation and Quality in the Wikipedia. In *Proceedings of the International Symposium on Wikis and Open Collaboration*, pages 157–164, Montreal, Canada, 2007.

[31] C. Wilson, B. Boe, A. Sala, K. P. Puttaswamy, and B. Y. Zhao. User interactions in social networks and their implications. In *Proceedings of the 4th ACM European Conference on Computer Systems*, pages 205–218, Nuremberg, Germany, 2009.

[32] G. Wu, M. Harrigan, and P. Cunningham. Characterizing wikipedia pages using edit network motif profiles. In *Proceedings of the 3rd International Workshop on Search and Mining User-generated Contents*, pages 45–52, Glasgow, Scotland, UK, 2011.

[33] D. Yang, A. Halfaker, R. Kraut, and E. Hovy. Edit Categories and Editor Role Identification in Wikipedia. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*, pages 1295–1299, Portorož, Slovenia, 2016.

[34] D. Yang, A. Halfaker, R. Kraut, and E. Hovy. Who Did What: Editor Role Identification in Wikipedia. In *Proceedings of the 10th international AAAI Conference on Web and Social Media*, pages 446–455, Cologne, Germany, 2016.