

# Lost in Translation: Contexts, Computing, Disputing on Wikipedia

Pasko Bilic<sup>1</sup> and Luka Bulian<sup>2</sup>

<sup>1</sup> Department of Culture and Communication, Institute for Development and International Relations, Zagreb

<sup>2</sup> Department of Sociology, University of Zagreb

## Abstract

Wikipedia is an open collaboration, global, multilingual project. Its guidelines and policies direct the collaboration process into a vision of objective and neutral encyclopedic knowledge. However, coherence of that knowledge, and the outcomes of the collaborative process on the same topic, can sometimes vary dramatically across different languages. Our goal was to explore what these differences are, and to see how they are contextualized in a case of a contested and conflictive topic. The empirical focus was on the Republic of Kosovo, a recently formed country in Southeast Europe still seeking full international recognition. The study explores the social, cultural and political tensions through following the contextualization of this topic in three different Wikipedia communities: Serbian, Croatian and English. A constructivist (Charmaz, 1998) and substantive grounded theory of the process was created by following a two-step coding process. Three coders were active in different stages of the process. Discussions and comparisons of emergent codes, within and between three different communities, were conducted regularly. The core concept of our theory was neutrality dispute. It is based on four aspects: identities and viewpoints, their input into the process of content editing, relations between the editors, and the process of conflict management. The main drivers of conflict and/or consensus, within and across languages were different types of group identifications in relation to the topic of Kosovo and Wikipedia in general. Wiki software and Wikipedia's rules help in managing multiple conflicts, although the political and cultural contentiousness of the topic existing in the "offline" context was also reproduced in the collaborative process.

**Keywords:** Wikipedia, identity, language, ideology, context, conflict, grounded theory

**Citation:** Bilic, P., & Bulian, L. (2014). Lost in Translation: Contexts, Computing, Disputing on Wikipedia. In *iConference 2014 Proceedings* (p. 32-44). doi:10.9776/14027

**Copyright:** Copyright is held by the authors.

**Acknowledgements:** Our gratitude goes to the Wikimedia Foundation's Research Committee for granting us with the approval to recruit and interview Wikipedians. We also thank Tea Vidović from the Center for Peace studies in Zagreb for her assistance in the first round of coding. Senada Šelo-Šabić, Saša Čvrljak and Iva Kornfein from the Institute for Development and International Relations provided crucial information about the historical and political context of Kosovo. Their help is much appreciated. We also thank two anonymous reviewers for providing valuable comments and guidelines for improving this paper.

**Contact:** pasko@irmo.hr, luka.bulian@gmail.com

## 1 Introduction

Wikipedia is a complex socio-technical system based on specific collaborative software and a set of rules, policies and guidelines for producing content. It runs on *MediaWiki* software which is a modified version of the original wiki software developed in 1995 to enable user-friendly, transparent collaboration and online database management. The normative side of Wikipedia was developed to ensure a clear set of rules for producing reliable encyclopedic content. The core principle is the neutral point of view (NPOV) which should present all sides without bias within a given article. With the growth of Wikipedia, specific rules and policies have multiplied making content management on bigger language versions an increasingly complex affair (Kittur, Suh, Pendelton and Chi, 2007).

Wikipedia consistently ranks among top ten sites on the web and is offered in almost three hundred different languages. However, Wikipedia's neutral and objectivist knowledge is not always easy to secure,

especially in light of the existing language and cultural diversity through which knowledge is differently contextualized. Instead of focusing on the technical and normative side of content production, we will explore the social and cultural dimensions of collaboration on a conflictive topic across three different language versions. The common technical and basic normative context of Wikipedia makes cross-language comparisons justified. Our goal is to explore what the differences are, and how these differences are contextualized. The empirical focus is on the recently formed Republic of Kosovo in Southeast Europe. As a highly contested, but notable, topic it sparked heated regional and international political debates, gained significant media coverage, and motivated editors across language versions to contribute with related articles on Wikipedia. Its conflictive nature makes opposing points of view apparent and makes an objectivist and neutral content production hard to reach. Due to the previous integration of the partially internationally recognized Republic as a province within the borders of Serbia, the topic is particularly contested on Serbian Wikipedia.

We employ a qualitative, explorative approach and compare editing dynamics, norms and values during the process of producing content about the Republic of Kosovo on Serbian, Croatian and English Wikipedia. Gathered data was analyzed using the constructivist grounded theory approach (Charmaz, 1998) in an iterative two-step coding process within and between three different interaction contexts and language versions of Wikipedia. Due to the sensitivity of the topic, three coders were active in different stages of the process to ensure a higher degree of interpretative flexibility. The goal was to create a substantive grounded theory (Glaser and Strauss, 1967) explaining the editing process of this specific and problematic topic in several selected languages. Our research shows that representing all sides of the story in an article is extremely complex, especially in a situation that has historical burdens, emotional tensions, and rapidly changes and evolves. The main drivers of conflict and/or consensus within and across languages were different types of self- and group identifications in relation to the topic of Kosovo. Wiki software and Wikipedia's rules help in managing the conflicts, but the social, cultural and normative aspects of these identities make it complicated. Nonetheless, Wikipedia offers an online mediation forum where anonymous contributors with diverging positions directly meet and negotiate under a common technological and normative context.

## 2 Wikipedia: global, multi-language, knowledge-building environment

Launched on 15 January 2001, Wikipedia is now Ranked 7<sup>th</sup> on Alexa's top sites list.<sup>1</sup> While the quality and accuracy of the articles is sometimes heavily debated due to Wikipedia's "anyone can edit" approach, two intrinsic control mechanisms regulate the editing process, namely the five pillars<sup>2</sup> and multiple policies and guidelines. The five pillars explain what Wikipedia is and what it stands for (a free encyclopedia with no firm rules, written from a neutral point of view by the public, for the public by editors who, above all, are to treat each other with respect and civility). Policies and guidelines expand on the main pillars; give guidelines on the content of articles, conduct, deletion, enforcement of various standards, legal considerations and remedies, etc. The core content production policies - verifiability and a neutral point of view, with no original research - demand an article to be written in a way in which it represents "fairly, proportionately, and... without bias, all of the views that have been published by reliable sources on a topic."<sup>3</sup>

---

1 Retrieved from <http://www.alexa.com/topsites>

2 Retrieved from [http://en.wikipedia.org/wiki/Wikipedia:Five\\_pillars](http://en.wikipedia.org/wiki/Wikipedia:Five_pillars)

3 Retrieved from [http://en.wikipedia.org/wiki/Wikipedia:Neutral\\_point\\_of\\_view](http://en.wikipedia.org/wiki/Wikipedia:Neutral_point_of_view)

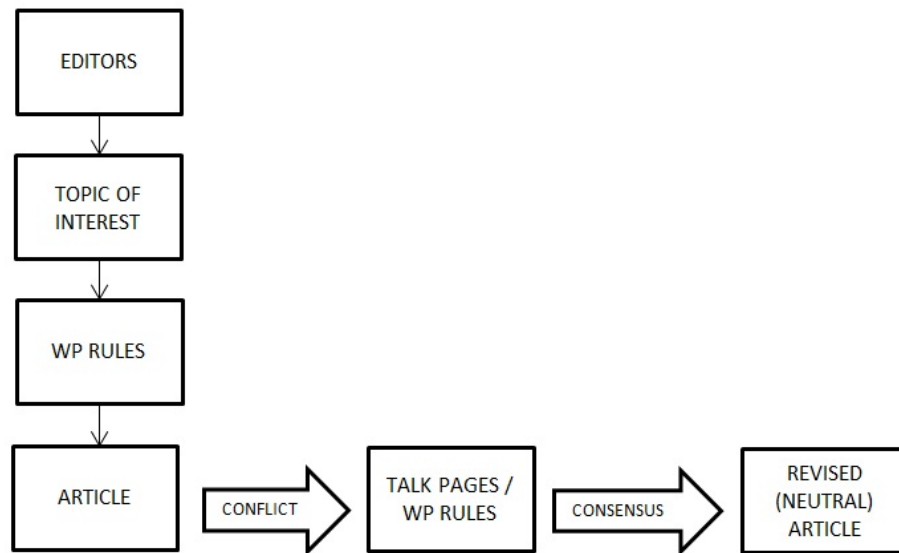


Figure 1: Standard content production model (objectivist)

By closely adhering to these rules and policies, content production would likely take the form presented in the diagram above. There is an editor, or a group of editors, interested in writing about a certain topic. By following the basic rules they create, or revise, an article. If there are disputes and conflicts, they are discussed and resolved on the talk pages. Consensus is reached and a revised article created. This process should, in principle, work the same way in all language versions.

### 3 Related research

Social interaction on Wikipedia and the organization of the work process have seen research focusing on the consensual aspects of community identity, and conflictive and disruptive behavior of different kinds. Bryant, Forte and Bruckman (2005) showed how long-term and frequent exposure to the editing process changes editors' perceptions of Wikipedia. They start identifying themselves strongly with the project, and start thinking of it as a community of co-authors with distinct roles and talents. Pentzold (2011) similarly showed that Wikipedia is an "ethos-action community". In other words, community membership and its boundaries are defined by a set of standards regarding Wikipedia's purpose, norms, values, and valid actions.

Viegas, Wattenberg and Dave (2004) used visualization software to study article revision history and have found different patterns of cooperation and conflict. Conflicts were evident as series of reverts in article revision history. Other forms of disruptive behavior include, for example, trolling as repetitive, intentional, and harmful violations of Wikipedia's policies (Schahaf and Hara, 2010). However, as Kittur et al. (2007) showed, conflict should not only be seen in a negative context, since it also offers positive benefits such as resolving disagreements, establishing consensus, clarifying issues and strengthening common values.

Some studies have started probing into Wikipedia's language diversity trying to compare communities according to different criteria. Pfeil, Zaphiris and Ang (2006) conducted research on the connection between collaborative patterns on Wikipedia and cultural and national backgrounds of editors on four language Wikipedias. The study shows that Wikipedia is not a culturally neutral space and that differences in behavior can be observed. Hecht and Gergle (2010) performed an analysis of 25 different language versions comparing the concepts included in each edition, and comparing the ways in which these concepts were described. The authors challenge the so-called "global consensus hypothesis" of objective and neutral knowledge and demonstrate significant cultural diversity across different languages on Wikipedia.

Yasseri, Spoorri, Graham and Kertész (2013) analyzed similarities and differences between controversial topics in 10 different language Wikipedias, simultaneously relating them to geographic locations. The authors conclude that different-language-Wikipedias demonstrate divergent social-spatial priorities, interests and preferences.

Quantitative, big data analyses have confirmed that the objective and neutral notion of knowledge is hard to reach, given the substantial language diversity of Wikipedia communities. However, while they do detect diversity, they cannot completely understand, or explain, the actual processes of managing, negotiating and establishing these differences. Culture is closely determined by language and, despite intensive globalization and transnational processes, still to an extent by territory. But culture is not simply a homogenous territorial and language-dependent concept. It is also dependent on different identities within a certain language and territory in a process of their constant discursive formation and negotiation.

## 4 Methods

Grounded theory (Glaser and Strauss, 1967; Strauss and Corbin, 1990; Charmaz, 1998) approach refers to a research process well adapted to explorative studies of human interaction in specific contexts. The focus is on inductively creating mid-range theories through systematic and iterative process of parallel data collection, coding and analysis. The creation of a so-called substantive theory was the goal of our research. In other words, the creation of a theory grounded in the empirical situation, and developed without forcing it into pre-conceived theoretical ideas (Glaser and Strauss, 1967). A completely open mind, without reliance on previous theoretical and methodological ideas or cultural assumptions, advocated by Glaser and Strauss (1967) is never entirely possible. Nonetheless, we took great care to ensure our emergent codes and categories were not forced into existing theoretical discourses or assumptions.

The initial fieldwork site was the Serbian Wikipedia and the *Kosovo and Metohija* article, the only existing article on the topic of Kosovo in Serbian language at the start of our research. Since the Republic of Kosovo was previously a province within Serbia, the editing process in this language community was selected as the first research site. The main data sources throughout the study were Republic of Kosovo related article and user talk pages<sup>4</sup> as a form of “non-reactive data” (Janetzko, 2008). These documents provide detailed records of collaborative interaction as well as discussions on problems and issues that arise throughout the editing process. Talk pages also provide a transparent record of the work process in a longitudinal manner since every statement is signed and time-stamped by the editor. We have also issued calls for interview participation to individual users in three different communities, have conducted one interview, and engaged in e-mail or talk page discussions on individual talk pages with several editors during the first and second phase of the coding process. Although we expected to interview more respondents, even short discussions provided valuable insights into community values. For example, given our clearly stated Croatian background in the call for research participants and the use of Croatian language and Latin script (Croatian is closely related to Serbian as a South Slavic language), entering the Serbian community was met with suspicion by some editors. They have openly opposed our research calling it “inappropriate for their project” on one occasion, and “science fiction” in another. Despite detailed project description and careful observance of ethical standards for online research, there was clear resistance to our entrance to the community. These discussions also made us aware of the process of mutual construction of the research process between us as researchers and our respondents. It made us realize our own pre-conceptions and cultural identities in framing the research process (Charmaz, 1998). The exposure to the underlying importance of identities within these communities had set the course for our research and marked the emergence of significant categories in the ongoing coding process.

---

<sup>4</sup> See Annex for a list of documents.

We used a two-step coding process (Saldaña, 2009). In the first cycle we used *in vivo*, process and initial coding. *In vivo* was used to record particularly salient wording, "process" to record the editing dynamic, and "initial" for other data that was difficult to label. The gathered codes were then tested, purified, compared and organized through second-cycle focused, axial and theoretical coding on new sampled documents and new interaction contexts. Three coders were used in the first-cycle coding on Serbian Wikipedia to keep an open minded approach and to ensure a high level of interpretative flexibility. After each analysis cycle meetings and discussions were organized to exchange the codes and memos, and to discuss emerging patterns. Theoretical sampling of documents was performed through following the analytic leads such as references to specific talk pages, related topics or the creation of new topically-related articles. For example, in the case of the English Wikipedia there were constant references to the Arbitration Committee's decision to impose a "one revert per week" rule to deter disruptive edits. Hence the document containing the original Committee decision was sampled to saturate the emergent "WP rules" category as part of the conflict management process. This revealed the lack of similar third-party mediation possibilities on smaller language Wikipedias. Also, the Serbian community created the Republic of Kosovo article shortly after the start of our research making it necessary to sample these documents as well.<sup>5</sup> Upon finishing coding of the Serbian Wikipedia, we expanded to include the Croatian and English versions.<sup>6</sup> This provided the opportunity to test the existing codes from the Serbian sample in new interaction contexts. Theoretical saturation was quickly reported by both coders on Croatian and English data. However, certain procedural and value-related differences were evident, coded and categorized into new elements of the emerging theory.

## 5 Results

### 5.1 Political context of the Republic of Kosovo

The Republic of Kosovo is a partly internationally recognized country in Southeast Europe, previously the Autonomous Province of Kosovo and Metohija within Serbia. The majority of the population consists of ethnic Albanians, except in northern municipalities where a Serbian majority population lives. Starting in the late 1980s nationalist tensions created a highly volatile situation with Serbia revoking the autonomous province status. This caused resistance, riots and an insurgency against Serbian repressive measures. Insurgent and massive counter-insurgent military campaigns from Serbia ensued. With international mediation bringing no results in calming the situation, a three-month NATO bombing campaign forced Serbia to withdraw its military and police presence in 1999. Following a UN Security council resolution, Kosovo was placed under the UN Interim administration. The negotiation process between Serbia and Kosovo ended without agreement and in 2008 Kosovo declared independence. A European Union-led law-and-order mission was given the task of improving the rule of law, while a NATO-led peacekeeping force was given the task of providing a secure environment. In late 2008 Serbia challenged the legality of the independence declaration before the International Court of Justice. In 2010 the Court released the opinion that the declaration did not violate general principles of international law. Nonetheless, Serbia still disputes its independence and dialogue between the two countries has brought very limited success. An EU mediated agreement was reached in April 2013 that, among other agreements, enabled Serbia to get a date for EU accession talks, and for Kosovo to gain more control over the Northern Kosovo municipalities. Despite the fact that Kosovo joined various international organizations and gained recognition from a large number of countries including the United States, its independence is also disputed by, among others, five EU countries (Greece, Cyprus, Slovakia, Romania and Spain) and Russia which has a strong leeway in blocking its UN membership due to its veto power in the UN Security Council.

<sup>5</sup> See table 1 for more details on article creation timelines.

<sup>6</sup> The sampling of other languages such as the Albanian (Albanians being a dominant ethnic community in Kosovo) was not possible given our unfamiliarity with that language.

Figure 2: Kosovo in Southeast Europe<sup>7</sup>

## 5.2 Multiple contextualization on Wikipedia

Due to the contested nature of the topic it was contextualized in different ways in different languages. The table below displays the article creation timelines and different articles under which Kosovo is discussed. The Serbian language Wikipedia contained information about the Republic in the *Kosovo and Metohija* article until recently when a separate *Republic of Kosovo* article was created. The Croatian language contains related information under the *Kosovo* article, and the English version has two articles (*Autonomous Province of Kosovo and Metohija* and *Republic of Kosovo*) that were separated from the main *Kosovo* article after the declaration of independence.

Article title	Language version and article creation date		
	<i>sr.wikipedia.org</i>	<i>hr.wikipedia.org</i>	<i>en.wikipedia.org</i>
Kosovo	-	8/5/2005*	15/12/2001 <sup>8</sup>
Kosovo and Metohija	6/9/2004*	-	-
Autonomous province of Kosovo and Metohija	-	-	21/2/2008*
Republic of Kosovo	29/4/2013*	-	25/7/2009*

Table 1: Article creation timelines (\* signifies the analyzed articles)

Our study resulted in a total of 439 codes<sup>9</sup> with the core concept being neutrality under heavy dispute. Neutrality is dependent on four core elements: identities and viewpoints, their implications for the process of content editing, consequent relations between editors and types of conflict management. These four elements are grounded in 21 categories, and 55 sub-categories. The diagram below displays the relations between the core elements of the concept and the relations between those categories. The model is common to all three language versions, although variety in the intensity of existing identities and viewpoints can be discerned with consequences on the nature of disputes and types of conflict management. We provide explanations and examples in the following sections.

<sup>7</sup> Retrieved from [http://en.wikipedia.org/wiki/File:Kosovo\\_in\\_its\\_region.svg](http://en.wikipedia.org/wiki/File:Kosovo_in_its_region.svg)

<sup>8</sup> Kosovo article on English Wikipedia contains several thousand discussion pages. Due to the explorative character of this study, including limited research resources available, we were not able to fully analyze this talk page.

<sup>9</sup> The total number of codes is substantial, however, the majority of them relate to very subtle differences within categories and sub-categories. In the planning stage of our research we opted for dense coding in relatively small-sized documents.

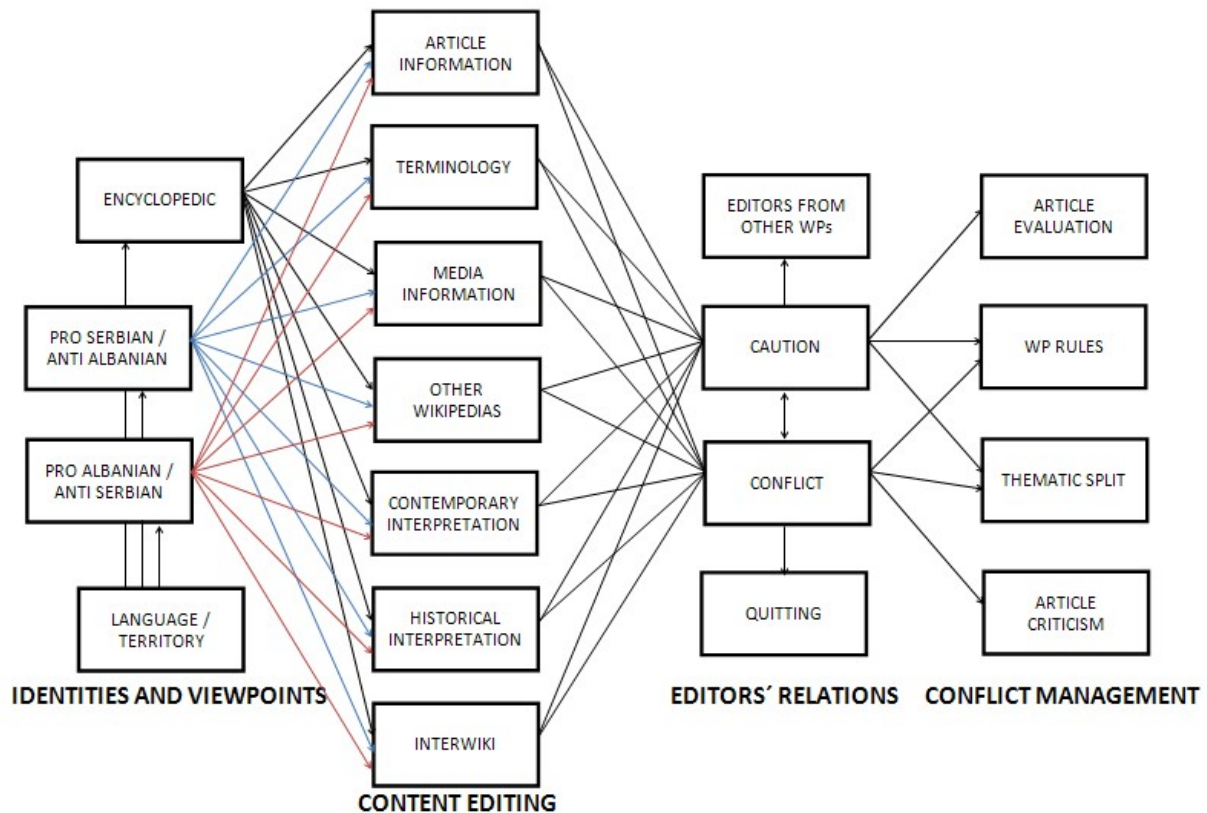


Figure 3: Conflictive content production model

### 5.2.1 Identities and viewpoints

Identities are explicit values and statements relating to self and group identification of a group of editors within Wikipedia in relation to the topic at hand. The dominant identity on all three languages was what we called the encyclopedic identity. It is closely related to the existing rules and policies of Wikipedia, but does not always address them explicitly during the collaboration process. This identity is evident within a group of editors who take the basic Wikipedia's policies, guidelines and goals as a form of their self-identification and who direct the editing process, manage relations and engage in interaction with other editors accordingly. It is a set of beliefs that Wikipedia should be careful when following current events and opinionated media reports; that objectivity is always achievable and highly desirable; that Wikipedia should always follow only the most relevant sources; that Wikipedia educates and informs people; and that Wikipedia is an international project based on clear rules. As one editor put it, explaining the objectivity principle: "It would be ideal if from the edits and comments ... one could not determine the political standpoint or nationality of the editor" (Razgovor: Kosovo).

Language and territorial identity is the second type of identity. It is not entirely coherent or homogenous within the analyzed language communities since it relates with different dynamics to the encyclopedic identity, and to differences in viewpoints towards Kosovo. For example, on both the Serbian and Croatian community, a clear separation between the language and the official state position towards Kosovo was explicit.<sup>10</sup> In that sense this identity type becomes closely related to the encyclopedic identity. However, it can also be used to legitimate the official state position. For example, in the Serbian Wikipedia: "This is SERBIAN Wikipedia. We have a right to our own opinion" (Razgovor: Kosovo i Metohija). It can

<sup>10</sup> The Republic of Croatia officially recognized the independence of the Republic of Kosovo.

also be used to demarcate the group boundaries from “outsiders” talking in different languages. For example, in the Croatian Wikipedia, certain editors were opposed to the Serbian grammar style used by a Serbian editor. Or in the English Wikipedia: “Your standard of English is very poor” (Talk: Republic of Kosovo).

These tensions get amplified by explicitly taking one side and simultaneously excluding the other in relation to Kosovo’s independence. They are not typical for one language community either. Nonetheless, the strongest pro-Serbian viewpoints can be found on the Serbian Wikipedia. They stem from an emotional response to the territorial loss (“painful topic”, “delicate subject” (Razgovor: Kosovo i Metohija)) and the perception of the international community as unjust towards Serbia (“Serbs are always to blame” (Razgovor: Republika Kosovo)). Simultaneously the independence is not accepted as legitimate and these positions sometimes turn into anti-Albanian attitudes and direct ideological and nationalist oppositions. They question the origin of the Albanian ethnic community and its religious orientation, and criticize the Albanian Wikipedia: “Unchecked, they can write whatever they want” (Razgovor: Kosovo i Metohija). The pro-Albanian viewpoints are weak in all three communities.<sup>11</sup> In the Croatian community they can be traced in connection to historical interpretations and the acceptance of the legitimacy of independence (“Kosovo was always inhabited by Albanians” (Razgovor: Kosovo)) or sometimes in its closely-related anti-Serbian and nationalist viewpoints. One editor made a sarcastic remark about the other: “You have a predictably anti-Serbian attitude and if 3 Martians landed in the Serbian Woodlands and proclaimed their state, you would grant them the right to do so” (Razgovor: Kosovo). These views, however, are not dominant but are in constant negotiation and struggle with the encyclopedic values and positions to impose certain interpretations on both local communities. The English Wikipedia seems to offer a forum for both Pro-Serbian and Pro-Albanian viewpoints making it difficult to negotiate a middle path between all of the existing identities and viewpoints.

### 5.2.2 Editing the article content

Actions based on previously defined identities and viewpoints lead to interpretations of different aspects of content editing and article structuring. This part of the process gets fragmented into a myriad of polarized debates. The issues often blend from one into the other. The most salient in all communities is the problem of a balanced introductory sentence and paragraph, followed by the organization of infoboxes and the remaining topical paragraphs. Article information often relates to detailed fact checking (list of states that have recognized the independence of Kosovo, demographic information, GDP, etc.) and using appropriate sources to fill in the blanks. Terminology relates to the way in which sentences and information is presented and described. Depending on the article context, local communities were focused on clarifying the status of the Republic as simply a territory, an autonomous province, or the Republic every time it was mentioned within the article. The English community faced a similar issue: “[t]he article is already crammed full of ‘partly recognized’ and ‘unilateral’ and so on, at every place in a sentence where somebody can cram in a caveat” (Talk: Republic of Kosovo).

Media and other information sources are rarely commented but can also be used to support or dispute certain claims. A significant amount of coordination work between language versions of Wikipedia existed, whether looked upon critically or positively. For example, the Serbian community was swinging between open criticisms towards a perceived bias in the English articles, towards acceptance of the thematic split between the articles: “The English solution is perfect” (Razgovor: Kosovo i Metohija). On the Croatian community other wikis were consulted to determine the frequency of use of the term Republic of Kosovo.

---

<sup>11</sup> It is noteworthy that the Republic of Kosovo’s Ministry of Foreign Affairs, in cooperation with several international organizations, organized the Wiki Academy Kosovo in 2013 to “[i]mprove the quality and quantity of online content on Kosovo to better represent Kosovo to the world.” The Academy included monetary prizes for best articles and photos representing Kosovo. Retrieved from <http://wikiacademykosovo.org/>



The English community also sought differences in contextualizing the topic, finding similarities between the Serbian and Russian Wikipedia.

The area of contemporary and historical interpretations of Kosovo-related topics is where ideological struggles are most evident. Contemporary interpretations relate to criticisms of international law as being “useless” or the declaration of independence as being a “shady area”. The international community is seen as using Kosovo as its satellite, denying Serbian constitutional demands for Kosovo. Historical interpretations relate to the ethnic settlement of the area by Serbs and Albanians and demographic changes of the area through history. Finally, links between Wikipedias (interwiki links), intended to directly link the same-topic articles between languages were disruptively edited in the English Wikipedia. Since the main Kosovo article was split into the Republic of Kosovo article, the creation of the article was perceived as a sort of “wiki-recognition” of the existence of that state. Links to other Wikipedias having similar Republic of Kosovo articles were deliberately hidden (Wikipedia: Arbitration/Requests/Enforcement).

### 5.2.3 Relations between the editors

Instances of editors from other Wikipedias doing cross wiki commenting on local communities are mostly episodic. However, editing the English Wikipedia is seen as an opportunity for giving greater visibility to one’s positions, whether encyclopedic or political. A number of editors with user accounts on local versions were active in the editing process on the English Wikipedia.

Exposure to the complex editing process leads to different relations between editors. In general, editors who display encyclopedic identity will seek more cautious relations with other editors, taking into account their positions and trying to find a common ground. Caution relates to giving thought to ones edits, taking time with reverts, thinking through the thematic split of related articles, being polite towards new editors, etc. For example one editor stated: “[t]his is an encyclopedia so we cannot play dumb and pretend something [Republic of Kosovo] does not exist” (Vikipedija: Članci za brisanje).

Conflicts arise from directly opposing positions between editors, breaking previous rules, agreements and neutrality principles. Conflicts and disruptive edits lead to a general lack of dialogue and open labeling of other editors through sarcasm, cynicism, open political opposition or malicious personal attacks. For example one editor expressed his frustration in the following way: “I’ve had it with fighting windmills, getting no support and being insulted...” (Razgovor: Kosovo). This poses the greatest challenge for Wikipedia as a project since it may lead to quitting from article editing or the project in general. It was often expressed as a feeling of frustration with the lack of consensus, incessant pushing of particular views, and repetition of similar non-constructive arguments: “I am at a loss as to how to move this debate beyond the “is not!” - “is too!” stage” (Talk: Republic of Kosovo).

### 5.2.4 Conflict management

Conflict management relates to all types of managing neutrality disputes and conflicts. It is performed through constant article evaluations or criticisms, adherence to rules and policies or alleviating conflicts by performing thematic splits between closely related topics. It is mostly done by administrators during neutrality or revert disputes. The measures include calls for discussions on talk pages or the introduction of different levels of article protection.

Caution in approaching the editors and the editing process leads to the positive evaluation of the article structure, its composing parts and overall existence: “We cannot ignore the existence of the Republic of Kosovo” (Razgovor: Republika Kosovo). Caution also leads to careful adherence to Wikipedia’s rules and policies in mediating the conflict and a carefully conducted thematic split. This was heavily debated in the English community where a split from the *Kosovo* article and the creation of *Autonomous province of Kosovo and Metohija* and *Republic of Kosovo* occurred. In the Serbian community the *Republic of Kosovo* article was split from the original *Kosovo and Metohija* article. However, it was heavily criticized as being a poor copy of the English versions: “Incorrect junk from the English wiki” (Razgovor: Republika Kosovo).

Conflictive relations between the editors will generally lead to negative article evaluation and article criticism, disapproval of rules and policies, criticism or blocking of the thematic article split, etc. Conflict management becomes difficult since the editors do not always perceive a discrepancy between Wikipedia's rules and their points of view. For example, an administrator on the English Wikipedia noted: He "...does not seem to realize that there is any Point-of View (POV) problem with his edits" (Wikipedia: Arbitration/Requests/Enforcement). A key difference between local communities and the English version in conflict management procedures is the existence of the Arbitration Committee in the English community which has the authority to impose binding solutions in cases of serious disputes. Due to repeated conflicts a number of warnings and bans regarding the Balkans related topics were issued by the Committee in the past several years (Wikipedia: Requests for Arbitration/Macedonia). Its enforcement decisions have generally lead to greater caution in the editing process although persistent problems can still lead to disruptive edits and quitting the Wikipedia project.

## 6 Broader theoretical relevance and future research

The relevance of identity in understanding online behavior is nothing new. In the past two decades it gained prominence as one of the key research areas in internet studies. However, surprisingly little attention has been given to the issue within Wikipedia research, apart from focusing on the core community identity. Our research shows that different community identities (encyclopedic, language, territorial) are not fixed, but flexible and negotiable. In other words, identity is relational and includes "boundary" work between definitions of "us" and "them" (Lamont, 2006). However, the intensity of this process also depends on the topical focus of the collaborative process. As Stryker and Burke (2000: 288) claim, behavior is goal directed and changes the situation in order to match the meanings perceived in the situation with meanings held in the standard. A mismatch between ones identity and the meanings perceived in the situation will result in negative emotion, while a decreasing discrepancy will result in positive emotion.

Similarly, Rössel and Collins (2006, p. 515) believe that "emotional energy" is the driving force of all "interaction rituals" with people seeking to maximize their own level of emotional energy. Past experiences get accumulated and direct individuals in seeking, or avoiding, certain interactions. Maximization of one's emotional energy can be done at the expense of the other, which means that different exclusionary tactics, pressures and patterns of dominant behavior can be exerted toward groups with an identity perceived to be different from one's own. This is the point where open conflicts and ideological clashes occur. As van Dijk states (1998), selected values serve as the basis for group self-identification. Dominant groups use the integration of values to legitimate their opposition, disagreement and resistance and to emphasize the polarization between "us" and "them" by showing "us" in a positive way, and "them" in a negative way.

The issues of identity work, emotional energy and ideological struggles were particularly salient in our research. However, they are by no means exclusive to these selected language versions and communities. In future research the presented substantive theory of neutrality disputes should be tested on other cross-language and conflictive topics on Wikipedia, including different language communities in the region and globally. These could improve our understanding of conflicts, ideologies and the role of cultural identities in online contexts and help us understand if, and how, online forums might facilitate and/or manage conflicts and clashes.

## 7 Conclusion

Social interaction is always messy and difficult to put into clearly separate categories. This is especially the case when highly problematic political issues are being discussed and negotiated in three different, but also related, online contexts. However, understanding the background processes of structuring online information is vital in contemporary networked societies, especially regarding Wikipedia as one of the most popular

global websites. After closely examining the collaborative process in this study certain patterns are discernible. We have identified social, cultural and political aspects that drive the editing process and distort the objective, neutral content production. The context of a language community plays a crucial role, but its identity is not entirely homogenous or coherent. Common to all three language communities was the existence of a group of editors with strong encyclopedic identity. However, exclusionary, nationalist identities put the objective notion of knowledge construction under strain. Also, communities are not closed entities, but relate to, and observe the editing process in other languages. A simple translation of article names to other languages might not refer to the entirely same topic. We have seen that the Republic of Kosovo was differently contextualized in different language communities and at different points in time. The identities and viewpoints shape the relations between editors and lead to either cautionary or conflictive editing. Conflict management includes constant article evaluations, article criticisms, article forking and adherence to existing rules and policies. Its success will depend on the intensity of included identities and viewpoints, and also on the size of the community and the availability of different conflict management mechanisms.

Overall, it seems that language and territory do not produce coherent and homogenous wiki communities, and hence do not produce homogenous and coherent knowledge. On the contrary, it appears that these Wikis re-produced political and cultural conflicts and diversity already existing in the "offline" contexts, adding to them the existence of purely online identities, such as the encyclopedic identity. The difference in comparison with the "offline world" is the possibility of direct, repeated confrontation between diverging positions in the negotiation process during collaborative creation of encyclopedic articles. In that sense online contexts such as Wikipedia provide mediation forums and environments for discussing difficult and problematic contemporary issues. Neither consensus nor conflicts are stable behavioral patterns on Wikipedia. Even small concessions of shortly lived consensus after long-term conflicts regarding minute details of a given topic add-up to the creation of diverging encyclopedic articles in different languages. These processes significantly alter the outlook of encyclopedic knowledge as represented in Wikipedia's articles. While related research has already emphasized the language diversity of Wikipedia's communities it provided limited explanation as to how and why these differences occur. A transversal approach to analyzing large amounts of quantitative data can distort the results since conflictive situations are prone to quick alterations over short periods of time. These dynamics can be better detected by taking note of the historical, cultural, political and other contextualizations, and by closely analyzing complex internal negotiations in a longitudinal manner. A given language should not be taken as an essential character trait of individual Wikipedia communities, nor should it be equated with geographic location or culture as a whole. We hope this study will provide further impetus for studying these differences and for shedding new light on the ways in which online and offline contexts interact and provide the capacity to change historical processes by offering alternative and novel communicative processes and negotiation forums. Whether Wikipedia's neutrality and objectivity principle can always transcend intense clashes and conflicts is yet to be studied and confirmed.

## 8 References

- Bryant, S. L., Forte, A. & Bruckman, A. (2005). Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia. *GROUP '05: Proceedings of the 2005 international ACM SIGGROUP conference on supporting group work*, 1-10. doi:10.1145/1099203.1099205
- Charmaz, K. (1998). Grounded Theory: Objectivist and Constructivist Methods. In: N. K. Denzin & Y. S. Lincoln (Eds.), *Strategies of Qualitative Inquiry* (pp. 509-535). Thousand Oaks, London, New Delhi: SAGE Publications.
- Corbin, J. & Strauss, A. (1990). Grounded Theory Research: Procedures, Canons, and Evaluative Criteria. *Qualitative Sociology*, 13(1): 3-21. doi:10.1007/BF00988593

- Glaser, B. G. & Strauss, A. L. (1967). *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago: Aldine.
- Hecht, B. & Gergle, D. (2010). The tower of Babel meets web 2.0: user-generated content and its applications in a multilingual context. *CHI 2010: Proceedings of the SIGCHI conference on human factors in computing systems*, 291-300. doi:10.1145/1753326.1753370
- Janetzko, D. (2008). Non-Reactive Data Collection on the Internet. In N. Fielding, R. M. Lee & G. Blank (Eds.), *The SAGE Handbook of Online Research Methods* (pp. 161-173). Los Angeles, London, New Delhi and Singapore: SAGE Publications.
- Kittur, A., Suh, B., Pendelton, B. A. & Chi, E.H. (2007). He says, she says: conflict and coordination on Wikipedia. *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, 453-462. doi:10.1145/1240624.1240698
- Lamont, M. (2006). Culture and identity. In J. H. Turner (Ed.), *Handbook of Sociological Theory* (pp. 171-186). New York: Springer.
- Pentzold, C. (2011). Imagining the Wikipedia community: what do Wikipedia authors mean when they write about their community? *New Media & Society* 13(5): 704-721. doi:10.1177/1461444810378364
- Pfeil, U., Zaphiris, P. & Ang, C. S. (2006). Cultural differences in collaborative authoring of Wikipedia. *Journal of Computer-Mediated Communication* 12(1): 88-113. doi: 10.1111/j.1083-6101.2006.00316.x
- Rösel, J. & Collins, R. (2006). Conflict theory and interaction rituals: the microfoundations of conflict theory. In J. H. Turner (Ed.), *Handbook of Sociological Theory* (pp. 509-532). New York: Springer.
- Saldaña, J. (2009). *The Coding Manual for Qualitative Researchers*. London: SAGE Publications.
- Schahaf, P. & Hara, N. (2010). Beyond vandalism: Wikipedia trolls. *Journal of Information Science* 36(3): 357-370. doi: 10.1177/0165551510365390
- Stryker, S. & Burke, P. J. (2000). The past, present, and future of an identity theory. *Social Psychology Quarterly* 63(4): 284-297.
- Suh, B., Convertino, G., Chi, E. H. & Piroli, P.L. (2009). The singularity is not near: slowing growth of Wikipedia. *Wikisym '09: Proceedings of the 5<sup>th</sup> international symposium on wikis and open collaboration*, article no. 8. doi:10.1145/1641309.1641322
- Van Dijk, T. (1998). *Ideology: A multi-disciplinary approach*. London, Thousand Oaks, New Delhi: SAGE Publications.
- Viégas, F. B., Wattenberg, M. & Dave, M. (2004). Studying cooperation and conflict between authors with *history flow* visualization. *CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems*, 575-582. doi:10.1145/985692.985765
- Yasseri, T., Spoerri, A., Graham, M. & Kertész, J. (2013). The most controversial topics in Wikipedia: A multilingual and geographical analysis. In: P. Fichman & N. Hara (Eds.), *Global Wikipedia: International and cross-cultural issues in online collaboration*. Scarecrow Press, 2014. Retrieved from: [http://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=2269392](http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2269392)

## 9 Table of Figures

Figure 1: Standard content production model (objectivist) .....	34
Figure 2: Kosovo in Southeast Europe .....	37
Figure 3: Conflictive content production model .....	38

## 10 Table of Tables

Table 1: Article creation timelines (\* signifies the analyzed articles)..... 37

## 11 Annex (Key documents)

### *Serbian Wikipedia*

- Razgovor: Kosovo i Metohija (arhiva 1, 2) [Talk: Kosovo and Metohija (archives 1, 2)], retrieved on 6 May 2013
- Razgovor: Republika Kosovo [Talk: Republic of Kosovo], retrieved on 14 May 2013
- Vikipedija: Članci za brisanje, arhiva 50 [Wikipedia: articles for deletion, archive 50], retrieved on 14 May 2013

### *Croatian Wikipedia*

- Razgovor: Kosovo [Talk: Kosovo], retrieved on 17 June 2013

### *English Wikipedia*

- Talk: Republic of Kosovo (archives 1,2), retrieved on 6 June 2013
- Wikipedia: Arbitration/Requests/Enforcement, retrieved on 2 July 2013
- Wikipedia: Requests for Arbitration/Macedonia, retrieved on 3 July 2013
- User talk: Ottomanist, retrieved on 4 July 2013
- Talk: Autonomous Province of Kosovo and Metohija, archive 1, retrieved on 4 July