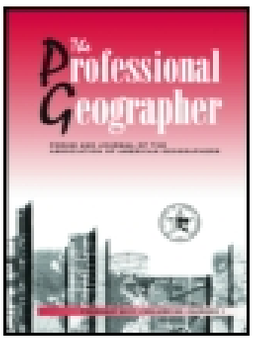

The Professional Geographer

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rtpg20

Edit Wars in a Contested Digital City: Mapping Wikipedia's Uneven Augmentations of Berlin

Cailean Osborne , Mark Graham & Martin Dittus

To cite this article: Cailean Osborne , Mark Graham & Martin Dittus (2020): Edit Wars in a Contested Digital City: Mapping Wikipedia's Uneven Augmentations of Berlin, The Professional Geographer, DOI: 10.1080/00330124.2020.1800493

To link to this article: https://doi.org/10.1080/00330124.2020.1800493

Published online: 11 Sep 2020.

Edit Wars in a Contested Digital City: Mapping Wikipedia's Uneven Augmentations of Berlin

Cailean Osborne
University of Oxford

Mark Graham and Martin Dittus
University of Oxford and Alan Turing Institute

Wikipedia is an information geography made up of millions of geotagged articles, which augment places with digital layers. These layers shape how people understand, move through, and interact with the world. As such, it is crucial to interrogate how places are augmented with digital layers and by whom. This article builds on the digital geography literature with a novel methodology that can be used to understand the digital representations of cities on Wikipedia, their provenance, as well as the scope and scale of "edit wars" that have shaped such layers. This methodology is demonstrated through a case study on Berlin, Germany. Cartographic analysis of about 8,000 geotagged articles reveals that these articles cluster in the city center, whereas outer districts are largely unrepresented on Wikipedia. Ten articles have had "edit wars" and nine of these are about contentious historical periods. Cartographic analysis of the locations of the editors of these articles reveals that the edit wars were mostly waged outside of Germany, raising questions about whose voices prevail in the digital representation of local places. The article concludes with a discussion of the results and directions for future work on the development of the methodology and the analysis of further cities. **Key Words: computational social science, Geoweb, information geography, Internet geography, Wikipedia.**

Local places are augmented by layers of digital information that shape not just how people understand the world but how they move through it and interact with it. As such, it is of crucial importance to understand who has the power to shape those digital layers of place. The crowdsourced encyclopedia, Wikipedia, is one of the world's most used, and therefore most important, digital augmentations of place. The digital layers on Wikipedia are produced through its open authorship model, which allows, in theory, any Internet user to create or edit geotagged information. This model aims to democratize how knowledge is produced and how the world is represented. Prior research, however, has demonstrated that the information geography of Wikipedia—that is, the geography of geotagged articles—is highly uneven and clustered in developed countries (Graham et al. 2014), and much of the world is largely represented by foreign-language—in particular, English—articles (Dittus and Graham 2019).

Although the Web site's information geography has been studied on a global scale, to date there has been little attention on the digital representation of cities. This article builds on the digital geography literature with a novel methodology that can be used to understand both the digital spatial representations of whole cities as well as the provenance of these representations. With the application of an "edit war" detection algorithm to the publicly available edit histories of geotagged Wikipedia articles within the bounds of a city, the article offers a method for making visible previously unnoticed processes of contestation between editors who have sought to shape the contents and hence nature of such digital layers. The methodology is demonstrated through a case study on Berlin, Germany, a city with a deeply contested geopolitical history, where multiple recent cultural and political histories have left traces in public space: recent enough to still be visible, but long ago enough to be removed from the everyday experience of the place. This makes it an appropriate candidate for a case study to trace how contesting narratives shape or are augmented by digital layers on Wikipedia. We hope that future research can extend this work to other global cities.

To achieve its objectives, this article seeks to examine three targeted research questions about the information geography of Berlin. First, it examines the visibility and legibility of the city to readers of Wikipedia, asking (RQ1) how Berlin is augmented on Wikipedia. Second, it examines (RQ2) how contested these information geographies are. Finally, it asks (RQ3) who gets to contest those representations of Berlin. Overall, this article builds on existing literature with an intervention that shows that we need to, and can, interrogate contestations of the digital layers of local places. It does that with a case study of Berlin and a replicable methodology that can be used for analyzing who gets to shape the digital layers of our cities.

The Professional Geographer, 0(0) 2020, pages 1–11
Initial submission, November 2019; revised submission, March 2020; final acceptance, May 2020.

Related Work

Information Geographies

Cities are augmented by digital information, which reflects and shapes the cities in bidirectional ways. Internet users impose narratives on space when they produce and share geotagged content, from TripAdvisor reviews to Facebook check-ins (Zook and Graham 2007). These augmentations do not just reflect but also shape people's relationships to places as they become "layered, defined, and augmented by information that is more or less immutable" (Graham, De Sabbata, and Zook 2015, 89). Thus, digital information becomes part of a place and how it is and can be known to the world (Graham 2013; Graham, Zook, and Boulton 2013; Graham, De Sabbata, and Zook 2015).

Wikipedia plays an important role in the construction of "geographical imaginations of place," and the knowledge gained from Wikipedia articles has an "immense power to augment our spatial understanding and interactions" (Graham et al. 2014, 760). The world is unevenly represented on Wikipedia, though, where "some parts of the world are at the center of global voice and representation, and many others [are] invisible or unheard" (Graham et al. 2014, 760). For instance, there are more geotagged articles in The Netherlands than in Africa as a whole (Graham et al. 2014). What is more, large swaths of the world are not just underrepresented but their representation tends to be disproportionately accounted for by foreign languages; in particular, English, which has emerged as a single dominant language in the representation of much of the world on Wikipedia, including the representation of many non-English-speaking countries in Africa, Asia, and South America (Dittus and Graham 2019).

Wikipedia's uneven information geography reflects how space is always embedded within "power geometries," implicating social groups with different relationships to space (Massey 1994). This is not new: "Geographies of codified knowledge have always been characterised by core-periphery patterns" (Graham et al. 2014, 746). On Wikipedia, uneven geographies are in part produced by "self-focus biases," where "articles about places, people, and events where the language of the edition was spoken were more prominent than those in other regions" (Hale 2014, 99; Hecht and Gergle 2009, 2010a, 2010b). Uneven representations between editions are especially pronounced in articles about local places and events, often but not always written in local languages (Kim et al. 2016). As the former example illustrated, an important exception to this rule is that in many parts of the world, socioeconomic realities and digital divides constrain participation in Wikipedia editing (Dittus and Graham 2019).

There is a crucial distinction to be made between rivalrous and nonrivalrous digital augmentations of geography on the social Web. On the one hand, nonrivalrous digital augmentations can be understood as a type of geotag that can be made by one Internet user that does not prevent another from using the same geotag. With the algorithmic logics of aggregation and indexing on Web sites, nonrivalrous augmentations can be searched, filtered, and discovered—hence, they are and can be visible and known—within constellations of complementary geotags. On the other hand, digital augmentations can be rivalrous when there can only be one augmentation for any given place. For instance, Wikipedia allows only one article per place per language edition. Within one article, editors can contribute to the online encyclopedia's information about this place. As the next section details, crowdsourcing knowledge about rivalrous augmentations of place is not always an uncontested process. Often, editors engage in edit wars to contest the contents of articles and ultimately have their views prevail in how such digital augmentations reflect and shape Wikipedia readers' understanding of real-world places.

Edit Wars on Wikipedia

The history of Wikipedia illustrates that crowdsourcing an encyclopedia is not an apolitical affair. Since the beginning of the Wikipedia project, the platform was designed to allow for consensus building between decentralized editors who contribute to Wikipedia with diverse motivations and represent "horizontal networks of interactive communication that connect local and global" (Castells 2007, 246). Editors do not always agree on facts; for instance, the country of origin of feta cheese is a controversial topic on Wikipedia: Editors who insist that feta cheese comes from Greece have systematically reverted the addition of other countries of origin, such as Macedonia, Turkey, and Bulgaria (Borra et al. 2014). Such conflicts occur when two or more editors repeatedly revert each other's contributions to articles—whether additions, deletions, or modifications of text. The Wikipedia community calls such conflicts edit wars and marks them through a variety of measures, including the three-revert rule, tagging disputes, and even listing "the lamest edit wars" (Yasseri et al. 2012). In individual articles, edit wars are identified in talk pages,[1] which often display—at times offensive—public discussions between editors and notices for "clean-up" (Yasseri et al. 2012). In some cases, discussions exceed the length of the associated article. Talk pages are used differently between editions, however, and are not always the best feature for spotting edit wars. For instance, editors in the English edition tend to use talk pages intensively to resolve disputes, whereas editors in the German edition do not use talk pages as much (Yasseri et al. 2012).

A more accurate way to detect edit wars is to inspect the edit history of an article, which is a public log of every edit made to an article complete with information about the editor (username if they are a registered editor or IP address if they are an anonymous editor), a timestamp of the edit, a link to the edited text, and the size of edit. If one sees that the same text has been reverted back and forth between two or more editors, this might indicate the presence of an edit war. Wikipedia contains millions of articles, though, each containing up to thousands of edits, which would make the manual investigation of edit wars a time-consuming task. This has led researchers to propose automated edit war detection methods. For instance, Kittur et al. (2007) and Vuong et al. (2008) enumerated how often dispute tags appear in articles. This approach, however, relies on editors having recognized and reported edit wars, leaving unnoticed edit wars indeed unnoticed by this approach. Sumi et al. (2011) developed an edit war detection classifier, which identifies contested editing patterns in article edit histories. Concretely, it looks for multiple pairs of authors who have repeatedly mutually reverted each other. The resulting contestation score per article increases if the warring editors are also highly active contributors to the article. Furthermore, the classifier separates edit wars from vandalism by assigning lesser weights to one-off mutual reverts. Studies using this classifier have shown that very few articles—less than 1 percent—have had edit wars (Sumi et al. 2011), and the most contested article categories are politics, geography, religion, and history by count of articles with edit wars (Yasseri et al. 2013).

Wikipedia: A Digital Memory Place

Editors have diverse motivations for augmenting and contesting articles on Wikipedia. One motivation is the expression and contestation of collective memory (Pentzold 2009; Ferron and Massa 2011, 2014; Borra et al. 2014; García-Gavilanes et al. 2016). As such, Wikipedia has been described as "a global memory place," where decentralized editors communicate and negotiate their views and memories of events and places (Pentzold 2009). The concept of a "memory place" was first articulated by Nora (1989), who distinguished living or lived memory environments—*milieux de mémoire*—and fixed and institutionalized memory places—*lieux de mémoire*—that can be "any significant entity, whether material or non-material in nature, which by dint of human will or the work of time has become a symbolic element of memorial heritage of any community" (Nora 1996, xvii). The institutionalization of memory is always selective and imbued with power dynamics, as the collective memory of some social groups is crystalized into national memory and outlasts other collective memories that are disqualified, erased, and forgotten (Halbwachs 1980).

The open authorship model of Wikipedia, in theory, presents a more democratic method of knowledge production and, indeed, memory work. It affords "greater diversity of voices, opinions, and narratives about any place," as anyone with an Internet connection can annotate a place with a geotagged article (Graham et al. 2014, 760). Wikipedia's architecture allows one to follow the "different observable steps of memory work as they evolve online" (Pentzold 2009, 267), as well as "look under the hood" into the contingent processes of contested memory in the talk pages and edit histories (Luyt 2016). For instance, publicly available edit histories document who has participated in the creation and contestation of articles—hence, where appropriate, in the creation and contestation of memory. The technical affordances of Wikipedia's architecture, however, do not dispel the persistence and in some cases the hegemony of prevailing discourses that shape what is or can be known about the world: "Even though Wikipedia consists of a massive cloud of geographic information about millions of events and places around the globe … it is characterized by uneven and clustered geographies: there is simply not a lot of content about much of the world" (Graham et al. 2014, 760; Borra et al. 2014).

Although the Wikipedia model might afford a more democratic process of spatial knowledge production and memory work, this does not necessarily manifest in practice, particularly when the underlying places are embedded in "power geometries" (Massey 1994), as "contemporary narratives as well as performances of subjectivity and authority are inscribed" to build identity and give structure to an immaterial past through space (Till 2005, 10). For instance, the decommunization of Berlin, as well as other central and eastern European cities, in the 1990s and the building of new memorials to remember the victims and crimes of the Third Reich reflected how social spaces in the city were being shaped according to new institutional needs, desires, and imaginaries of the new democratic German capital (Robinson 2004; Till 2005; Czepczyński 2008; Hirt 2012; Ferenčuhová and Gentile 2016).

Methods

Case Study Selection

Berlin was selected as a case study because the city has a deeply contested geopolitical history, where multiple recent cultural and political histories have left traces in public space: recent enough to still be visible but long ago enough to be removed from the everyday experience of the place. This makes it a good candidate for a case study to trace how such contesting historic narratives shape or are augmented by contemporary digital shadows.

Data Collection

Data for this analysis were obtained by retrieving all geotagged articles[2] on Wikipedia from a Wikimedia data dump[3] in February 2018. There were almost 13 million geotagged articles (14 percent of all 89 million articles on Wikipedia), which were filtered based on the condition of having coordinates within the greater area of Berlin. This resulted in a data set of 7,993 articles from seventy-three language editions. To map participation—that is, where in the world editors have edited these articles—we followed a tested methodology to collect and process geographic metadata of anonymous editors,[4] whose IP addresses[5] are made available by Wikipedia (Graham, Straumann, and Hogan 2015). Wikipedia not does not make the IP addresses of registered editors available and hence they are not considered in this study. The IP addresses allow us to determine the location of edits at country resolution. Although IP addresses do not perfectly correspond to distinct editors (for reasons explained in Graham, Straumann, and Hogan 2015), prior work shows that the resulting distribution is highly correlated with the overall global distribution of editors as a whole, including registered editors (Graham, Straumann, and Hogan 2015).

An edit war detection classifier was used to detect contestation between editors of articles. The classifier identifies instances in article edit histories where multiple pairs of authors have repeatedly mutually reverted each other's work (Sumi et al. 2011; Yasseri et al. 2012). Although this is merely one kind of contestation, it has the advantage that it is language independent and can easily be applied at a large scale. The classifier calculates scores per article based on the following specifications: (1) an edit war takes place when multiple pairs of editors keep undoing each other's edits; (2) the most active pair is ignored because it might simply constitute a personal feud between two editors; and (3) the intensity of an article's contestation is assumed to be higher in cases where the editors have been highly active contributors to the article. Articles with a score above 1,000 have had edit wars and articles with a score above 200 contain as much vandalism[6] as edit warring (Sumi et al. 2011; Yasseri et al. 2012).

Data Analysis

The analysis is structured in sequence of answering the three targeted research questions. First, the locations of geotagged articles were visualized on a map of Berlin with demarcations per city district (Figure 1). Red dots represent articles and bold black dots

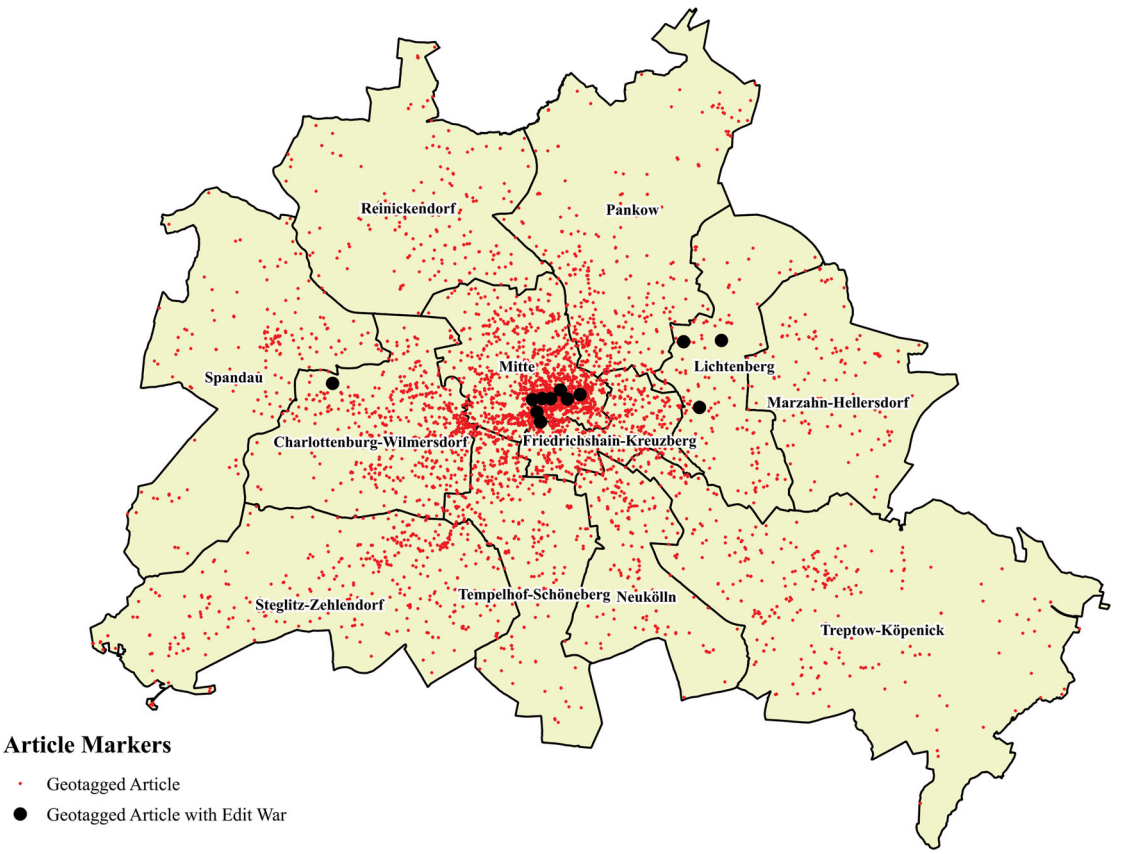

Figure 1 *Locations of geotagged articles in Berlin.*

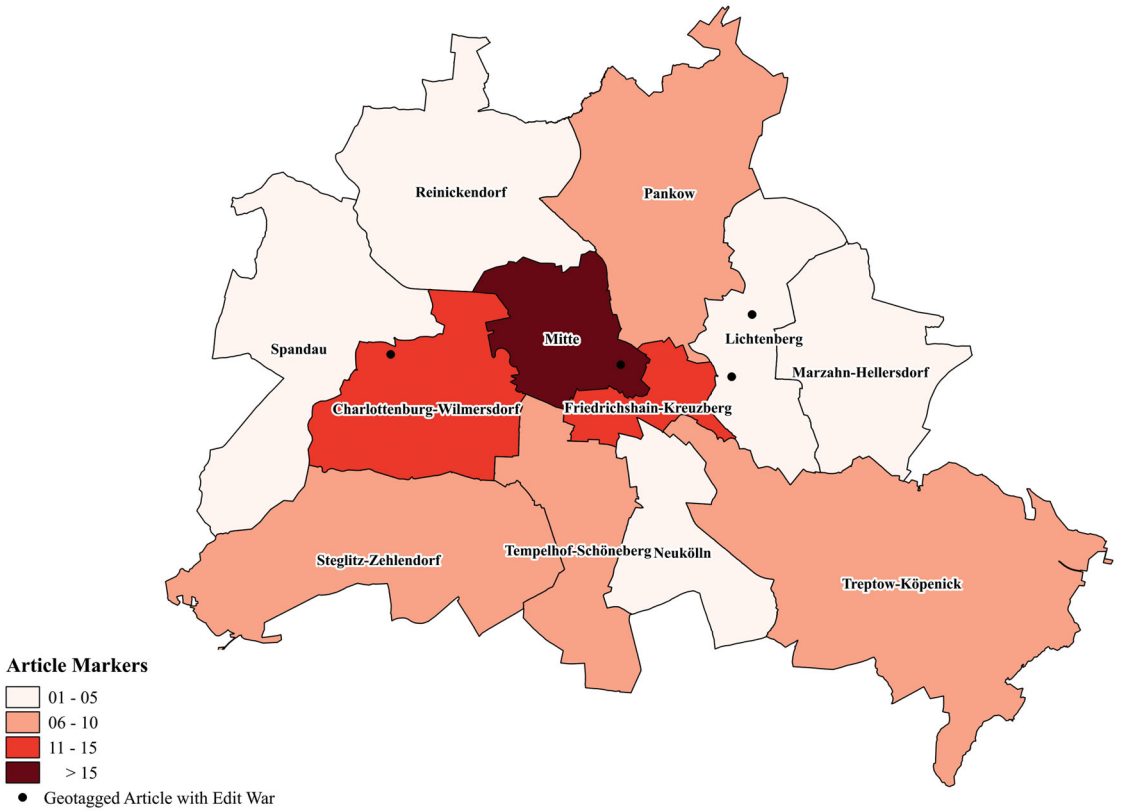

Figure 2 *Geotagged article density in German edition (N = 3,741). Note: Local population data from Amt für Statistik Berlin Brandenburg (2018).*

represent contested articles. The same data were visualized as choropleth maps (Figures 2 and 3), showing the density of articles in the German and English editions, respectively, normalized by each district's local population (in 10,000s).[7] These editions were chosen because they contain the most geotagged articles in Berlin (3,741 and 974, respectively). The districts were classified into four classes using the natural breaks (Jenks) classifier in QGIS, which is an optimization method that minimizes variation in each class but still separates outliers. The same classes were used for both choropleth maps of Berlin, despite their different article numbers (3,741 and 974, respectively), to allow for direct comparison. Following this, choropleth world maps were produced to visualize the global locations of anonymous editors who have edited all geotagged articles in Berlin (Figure 4) and subsequently only the contested articles (Figure 5). Edit counts per country were classified into four classes using the natural breaks classifier in QGIS. Whereas the Berlin maps were assigned the same classes due to roughly similar numbers of articles, the world maps were assigned their own classes due to significant differences in numbers of edits (164,721 and 16,563, respectively).

Results

Digital Representation of Berlin

Wikipedia's information geography of Berlin contains 7,993 geotagged articles, of which almost half are in the German edition (47 percent), followed by the English edition (12 percent), Italian edition (9 percent), French edition (5 percent), and Russian edition (3.5 percent). Some thirty-one editions contain between 10 and 1,000 articles, including Polish, Spanish, Hebrew, Chinese, and Turkish. Figure 1 shows that across Wikipedia's language editions, most articles cluster in Berlin's city center, Mitte, where many government buildings, museums, universities, and tourist attractions are based. Other central districts, such as Friedrichshain–Kreuzberg and Charlottenburg–Wilmersdorf, are moderately represented. The articles that are geotagged in the city center are not articles just about places and buildings but also about the city's history and its politicians. This reveals an important characteristic of Wikipedia's information geography: It is layered by a wide variety of social, political, and historical augmentations, which cast multifaceted

Figure 3 *Geotagged article density in English edition (N = 974). Note: Local population data from Amt für Statistik Berlin Brandenburg (2018).*

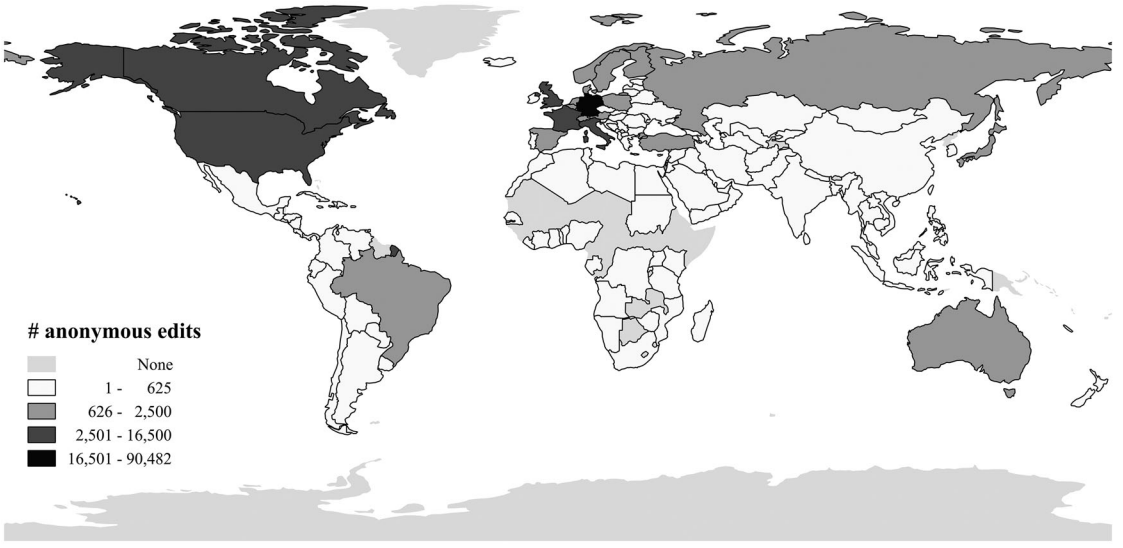

Figure 4 *Locations of edits to geotagged articles in Berlin.*

digital shadows on local places. The outer districts contain far fewer articles and are largely illegible, which shows that digital augmentations are clustered in already prominent and visible urban areas.

After normalizing for the local population, it becomes clear that outer districts have higher representation in the German edition than in other editions, suggesting local content production in these districts. Figures 2 and 3 compare the representation of the city

Figure 5 *Locations of edits to contested geotagged articles in Berlin.*

in the German edition (3,741 articles) and the English edition (974 articles). The maps illustrate that in the local German language, outer districts are represented in nontrivial ways, in particular Pankow in the north, Treptow-Köpenick in the southeast, and Steglitz–Zehlendorf and Tempelhof–Schöneberg in the southwest. Still, the outer eastern and western districts are poorly represented. The geography of Berlin in the English edition shows much more significant information clustering with comparably little representation of the outer districts.

With the exception of Mitte, all districts are more or less invisible. At the same time, the English edition contains most of the contested articles in the city, suggesting a vested interest among English-speaking editors to contest and shape the narratives about places (discussed more later). Meanwhile, the invisibility of outer districts in the English edition, the lingua franca of Wikipedia, as well as most language editions, has grave implications for the reproduction of what is already visible in the city as well as the invisibility and illegibility of usually marginalized parts of cities to readers of Wikipedia.

Contestation of Geotagged Articles

Only ten articles (0.1 percent of the global total) met the threshold of edit wars and are characterized by "very high controversy" (Sumi et al. 2011; Yasseri et al. 2012). Of these, three are in the German edition, six are in the English edition, and one is in the French edition (see Table 1). Figure 1 shows most of these articles are in Mitte, with the exception of three articles in Lichtenberg in the east. Under closer inspection, these three articles are about the GDR[8] in former East Berlin and two of these articles are in the German edition, reflecting local attention to this recent history. Seven of the ten articles are historical, including Battle for Berlin, Nazi Germany, Ministerium für Staatssicherheit (the Stasi), Berliner FC Dynamo (a GDR football club), and Gedenkstätte Hohenschönhausen–Berlin (a Stasi secret prison). After lowering the threshold to observe as much vandalism as edit wars (as described in the Methods section), some fourteen geotagged articles (0.2 percent of the global total) were identified as contested—hence, an additional four articles. Three of these articles are in the English edition, all of which are historical—German Empire, Gestapo, and Voßstraße (the location of Hitler's bunker)—and the remaining article is a German article about the football club Hertha BSC.

The results reveal that the contested augmentations of Berlin disproportionately relate to the city's twentieth-century history. It is worth noting that five of these articles are about the Third Reich, all of which are in the English edition. Another five are about the GDR, three in the German edition and two in the English edition. These results demonstrate that the representation of the city on Wikipedia is intimately connected to its history and that the English- and German-speaking editor communities have focused on different histories—the Third Reich and GDR, respectively—reflecting the prominence of certain histories in the collective memory of distinct language communities. Although the contestation of certain articles in editions whose languages are spoken in multiple countries or by nonnatives cannot be attributed to a distinct country or countries, finding such divergent flows of attention between language communities deserves careful attention regarding whose voices are heard in these contestations and ultimately prevail. Although it is

Table 1 *Contested geotagged articles in Berlin, Germany*

Edition	Article	Edit war score	No. edits
EN	Berlin	54,782	4,047
EN	Nazi Germany	32,768	3,750
EN	Battle of Berlin	23,454	1,120
DE	Berlin	7,440	1,275
DE	Ministerium für Staatssicherheit (Stasi)	6,450	762
EN	Schutzstaffel (SS)	5,040	1,715
EN	Berliner FC Dynamo	3,280	172
EN	Berlin Wall	2,880	3,279
FR	Angela Merkel	2,800	478
DE	Gedenkstätte Berlin-Hohenschönhausen	2,688	135
EN	Voßstraße	35	936
DE	Hertha BSC	1,355	390
EN	German Empire	1,257	360
EN	Gestapo	1,480	252

EN = English edition of Wikipedia, DE = German edition of Wikipedia, FR = French edition of Wikipedia.

beyond the scope of this study, future studies might benefit from close readings of contested articles and the examination of prevailing discourses that shape such articles' narratives about places, people, and events. In the next section, we offer one way of testing this question by providing a cartographic analysis of editors who have participated in these edit wars.

Participation in Edit Wars

Figure 4 shows that the geotagged articles were edited by anonymous editors on a global scale (some 164,721 anonymous edits). Most edits were made in Germany, followed by the United States, Canada, the United Kingdom, France, and Italy. The third largest grouping includes Spain, Scandinavia, parts of Eastern Europe, Russia, Turkey, Australia, and Brazil. The least editing occurred in Africa, Asia, and South America. Figure 5 shows that most contributors to edit wars were in the United States, which contrasts with Figure 4, which shows that Germany is where most edits are made in general. The United States is followed by Germany, the United Kingdom, Canada, France, and Australia in number of edits to contested articles, providing some evidence of an enduring fascination with Berlin's national-socialist history in English-speaking countries. Further cartographic analysis of the locations of editors of individual articles confirmed that the contested Third Reich articles in the English edition were mostly edited in Anglophone countries (although many were in Germany, too). Most editors who contested the German articles about the GDR and the French article about Angela Merkel were based in Germany and France, respectively.

Discussion

The empirical results illustrate that the urban information geography of Berlin on Wikipedia is highly uneven and clustered in the city center, leaving most outer districts illegible and invisible with the exception of the German edition. There is starkly uneven information clustering, with concentrations of articles in already prominent aspects of the city—that is, the city center—and its recent history are both highly represented and contested. With the exception of the German edition, outer districts are poorly represented in most language editions, indicating that the city's information geography on Wikipedia is marked by significant blank spaces, which make most of the city illegible and invisible to most readers of Wikipedia. These blank spaces matter, as increasing amounts of people become reliant on free sources of information to learn about the world. They set limits to how we can understand and interact with space. Moreover, this observation about Berlin is congruent with previous studies of Wikipedia, which have found that despite the crowdsourced encyclopedia's aims of reconfiguring information geographies by affording anyone with an Internet connection the opportunity to annotate any part of the world, geographies of codified knowledge continue to be uneven, regardless of platform affordances (Graham et al. 2014).

Relatively few articles have had edit wars, but they reflect a trend of editors engaging in heated debates about historical aspects, chiefly relating to the city's national-socialist and communist history. This finding suggests that geotagged articles have been shaped by processes of contested collective memory, where editors have engaged in spatial narrative building; thus, the articles are contested digital memory places. Like many informational augmentations, however, these digital memory places are not created evenly. The detection of edit wars in geotagged articles signals different contested memories about the history of Berlin on a global scale. Significantly, whereas most local augmentations were generated in the local German language, most contested articles were contested neither in the German language nor by editors based in Germany. The different geographies of participation in edit wars reflect divergent collective memories of the

city of Berlin: Aspects of the city's national-socialist period have received more active engagement by editors in many Anglophone countries, whereas controversial aspects of former communist East Berlin are mostly controversial among German-speaking editors. Although we cannot state that editors based in specific countries are necessarily citizens of these countries, finding a leading activity of editors in English-speaking countries—the United States, the United Kingdom, Canada, and Australia—is congruent with prior work on the enduring Anglo-American fascination with the Third Reich as well as unresolved tensions and sensitivities in Germany regarding the history of the GDR. To draw any further conclusions about how such cultural dynamics might be at stake one would have to examine the contents of the articles as well as the discussions in the talk pages more closely, which is beyond the scope of this analysis.

Uneven participation in edit wars could have grave implications for the ways in which local information geographies, including digital memory places, are represented and in turn how these representations reverberate back into the city in terms of how they might shape understandings of these places. The observation of uneven participation raises concerns for knowledge equity on Wikipedia as well as questions about whose knowledge and voices ultimately shape the digital augmentations of local places on Wikipedia. For instance, these findings illustrate that although Wikipedia is indeed a "global memory place," where memory places can be edited and contested in globally networked ways, it is not one that is evenly and necessarily locally produced. With these results, we hope to stimulate further research on the contested information geographies of cities on Wikipedia and how social, political, economic, and historical factors, among others, might play a part in these contestations.

Conclusion

This article has presented initial results on the unevenness of local information geographies on Wikipedia, with a case study on Berlin, Germany. The analysis demonstrates that the local information geography of Berlin is highly uneven, with most articles clustering in the city center and outer districts remaining unrepresented in most editions. It was observed that specifically the city's history is highly contested among editors and certain voices are more active than others in contesting Berlin's history on Wikipedia. Because information learned from articles inform understandings of places, people, and events, uneven representation and participation risk the reproduction of knowledge inequity by reflecting and reinforcing what is already made highly visible and widely known by other digital and analogue sources. Although this study examined articles at the level of their geotags and not their contents, future work might conduct closer readings of contents or talk pages to provide insights into the nature of edit wars. Emulating the methods presented here in linguistically contested cities such as Jerusalem, Cape Town, or Brussels will undoubtedly tell us much about how conflict in those cities manifests in their digital augmentations. Furthermore, whereas participation here was examined through the lens of anonymous edits, future research could seek to explore methods of accurately capturing geographic metadata of registered editors. Overall, as a first analysis of the information geography of a city on Wikipedia, this article seeks to stimulate more research in this direction in terms of both developing methodologies and examining further cities to make broader claims about Wikipedia's local information geographies. Mapping urban geographic inequalities on Wikipedia allows us to develop understandings of how they matter and ultimately develop strategies that can help in the design of more equitable futures. ■

Notes

[1] A talk page is a platform feature that allows editors to discuss improvements to an article.

[2] Geotagged articles are articles with geographic metadata. Geotags can be added to articles by any editor.

[3] The Wikimedia Foundation publishes whole databases (or "data dumps") of the language editions of Wikipedia on a weekly basis. Dumps include entire editions' articles and their complete edit histories.

[4] Wikipedia allows for registered and anonymous editing; the former entails editing by logged-in users whose edits can be traced back to their account and the latter entails editing by anonymous editors without accounts.

[5] An Internet Protocol (IP) address is a numerical label assigned to each device connected to a computer network, which functions for network identification and location addressing.

[6] On Wikipedia, vandalism is understood as an activity where an editor vandalizes the contents of an article, such as deleting or falsifying information in one-off edits.

[7] Local population data for Berlin's twelve districts were retrieved from Amt für Statistik Berlin Brandenburg (2018). "Statistischer Bericht: Einwohnerinnen und Einwohner im Land Berlin."

[8] The German Democratic Republic (GDR), often referred to as the former East Germany, was a real socialist regime between 1949 and 1990.

Funding

Martin Dittus and Mark Graham are funded by the Leverhulme Prize (PLP-2016-155), and Mark Graham is partially funded as a Turing Fellow

under Turing Award Number TU/B/000042. This work was also made possible through funding from the European Research Council under the European Union's Seventh Framework Programme for Research and Technological Development (FP/2007–2013)/ERC Grant Agreement No. 335716.

Literature Cited

Amt für Statistik Berlin Brandenburg. 2018. Statistischer Bericht: Ein wohnerinnen und Einwohner im Land Berlin [Statistical report: Residents in the country Berlin]. Accessed March 31, 2019. https://www.statistik-berlin-brandenburg.de/publikationen/stat_berichte/2019/SB_A01-05-00_2018h02_BE.pdf.

Borra, E., E. Weltevrede, P. Ciuccarelli, A. Kaltenbrunner, D. Laniado, G. Magni, and T. Venturini. 2014. Contropedia—The analysis and visualization of controversies in Wikipedia articles. In *OPENSYM: proceedings of the 10th International Symposium on Open Collaboration*, August 27–29, Berlin, Germany [G5]. New York: Association for Computing Machinery. https://doi.org/10.1145/2641580.2641622

Castells, M. 2007. Communication, power and counter-power in the network society. *International Journal of Communication* 1 (1):238–66.

Czepczyński, M. 2008. *Cultural landscapes of post-socialist cities: Representation of powers and needs*. Aldershot, UK: Ashgate.

Dittus, M., and M. Graham. 2019. Mapping Wikipedia's geolinguistic contours. *Digital Culture & Society* 5 (1): 147–64. doi: 10.14361/dcs-2019-0109.

Ferenčuhová, S., and M. Gentile. 2016. Introduction: Post-socialist cities and urban theory. *Eurasian Geography and Economics* 57 (4–5):483–96. doi: 10.1080/15387216.2016.1270615.

Ferron, M., and P. Massa. 2011. Collective memory building in Wikipedia: The case of North African uprisings. In *WikiSym 2011 Conference Proceedings—7th Annual International Symposium on Wikis and Open Collaboration*, 114–23.

Ferron, M., and P. Massa. 2014. Beyond the encyclopaedia: Collective memories in Wikipedia. *Memory Studies* 7 (1): 22–45. doi: 10.1177/1750698013490590.

García-Gavilanes, R., A. Mollgaard, M. Tsvetkova, and T. Yasseri. 2016. Memory remains: Understanding collective memory in the digital age. http://arxiv.org/abs/1609.02621.

Graham, M. 2013. Geography/Internet: Ethereal alternate dimensions of cyberspace or grounded augmented realities? *The Geographical Journal* 179 (2):177–82. doi: 10.1111/geoj.12009.

Graham, M., S. De Sabbata, and M. A. Zook. 2015. Towards a study of information geographies: (Im)mutable augmentations and a mapping of the geographies of information *Geo: Geography and Environment* 2 (1):88–105. doi: 10.1002/geo2.8.

Graham, M., B. Hogan, R. K. Straumann, and A. Medhat. 2014. Uneven geographies of user-generated information: Patterns of increasing informational poverty. *Annals of the Association of American Geographers* 104 (4): 746–64. doi: 10.1080/00045608.2014.910087.

Graham, M., R. Straumann, and B. Hogan. 2015. Digital divisions of labor and informational magnetism: Mapping participation in Wikipedia. *Annals of the Association of American Geographers* 105 (6):1158–78.

Graham, M., M. A. Zook, and A. Boulton. 2013. Augmented reality in the urban environment: Contested content and the duplicity of code. *Transactions of the Institute of British Geographers* 38 (3):464–79. doi: 10.1111/j.1475-5661.2012.00539.x.

Halbwachs, M. 1980. *The collective memory*. New York: Harper & Row.

Hale, S. 2014. Multilinguals and Wikipedia editing. In *WebSci 2014: Proceedings of the 2014 ACM Web Science Conference*, 99–108.

Hecht, B., and D. Gergle. 2009. Measuring self-focus bias in community-maintained knowledge repositories. In *Proceedings of the Fourth International Conference on Communities and Technologies, C&T '09*, 11–20. New York. doi: 10.1145/1556460.1556463.

Hecht, B., and D. Gergle. 2010a. On the "localness" of user-generated content. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work, CSCW*, 229–32.

Hecht, B., and D. Gergle. 2010b. The Tower of Babel meets Web 2.0: User-generated content and its applications in a multilingual context. In *Proceedings of the 28th International Conference on Human Factors in Computing Systems, CHI '10*, 291–300. New York.

Hirt, S. 2012. *Iron curtains*. Malden, MA: Wiley.

Kim, S., S. Park, S. Hale, S. Kim, J. Byun, and A. H. Oh. 2016. Understanding editing behaviors in multilingual Wikipedia. *PLoS ONE* 11 (5):e0155305. doi: 10.1371/journal.pone.0155305.

Kittur, A., B. Suh, B. Pendleton, and E. H. Chi. 2007. He says, she says: Conflict and coordination in Wikipedia. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, 453–62. New York.

Luyt, B. 2016. Wikipedia, collective memory, and the Vietnam War. *Journal of the Association for Information Science and Technology* 67 (8):1956–61. doi: 10.1002/asi.23518.

Massey, D. 1994. *From space, place and gender*. Minneapolis: University of Minnesota Press.

Nora, P. 1989. Between memory and history: Les lieux de mémoire. Representations 26 (26):7–24. doi: 10.2307/2928520.

Nora, P. 1996. *Realms of memory: Rethinking the French past: Vol. 1. Conflicts and divisions*. New York: Columbia University Press.

Pentzold, C. 2009. Fixing the floating gap: The online encyclopedia Wikipedia as a global memory place. *Memory Studies* 2 (2):255–72. doi: 10.1177/1750698008102055.

Robinson, J. 2004. In the tracks of comparative urbanism: Difference. *Urban Geography* 25 (8):709–23. doi: 10.2747/0272-3638.25.8.709.

Sumi, R., T. Yasseri, A. Rung, A. Kornai, and J. Kertész. 2011. Edit wars in Wikipedia. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, 724–27. doi: 10.1109/PASSAT/SocialCom.2011.47.

Till, K. E. 2005. *The new Berlin: Memory, politics, place*. Minneapolis: University of Minnesota Press.

Vuong, B.-Q., E. P. Lim, A. Sun, M.-T. Le, H. W. Lauw, and C. Kuiyu. 2008. On ranking controversies in Wikipedia: Models and evaluation. In *Proceedings of the international conference on Web search and web data mining '08*, 171–82. New York. doi: 10.1145/1341531.1341556.

Yasseri, T., A. Spoerri, M. Graham, and J. Kertész. 2013. The most controversial topics in Wikipedia: A multilingual and geographical analysis. http://arxiv.org/abs/1305.5566.

Yasseri, T., R. Sumi, A. Rung, A. Kornai, and J. Kertész. 2012. Dynamics of conflicts in Wikipedia. *PLoS ONE* 7 (6):e38869. doi: 10.1371/journal.pone.0038869.

Zook, M., and M. Graham. 2007. The creative reconstruction of the Internet: Google and the privatization of cyberspace and DigiPlace. *Geoforum* 38 (6):1322–43. doi: 10.1016/j.geoforum.2007.05.004.

CAILEAN OSBORNE is a former graduate student and research assistant at the Oxford Internet Institute, University of Oxford. His research interests are in the fields of computational social science and critical data studies; in particular, he is interested in the questions of participation and influence on online knowledge platforms as well as the development of computational methodologies for the social sciences. E-mail: caileanosborne@gmail.com.

MARK GRAHAM is the Professor of Internet Geography at the Oxford Internet Institute, a Fellow at the Alan Turing Institute, a Senior Research Fellow at Green Templeton College, a Research Affiliate in the University of Oxford's School of Geography and the Environment, a Research Associate at the Centre for Information Technology and National Development in Africa at the University of Cape Town, and a Visiting Researcher at Wissenschaftszentrum Berlin für Sozialforschung and Technische Universität Berlin. All of his publications are available at www.markgraham.space

MARTIN DITTUS is a digital geographer and data scientist with more than a decade of experience in social computing, mass participation platforms, digital geography, and big data. In his research, he applies quantitative and statistical methods, augmented with deep domain understanding, to analyze and visualize emerging online practices at large scale. E-mail: martin.dittus@oii.ox.ac.uk.