News coverage, digital activism, and geographical saliency: A case study of refugee camps and volunteered geographical information

Ron Mahabir[1,2]*, Arie Croitoru[1,2], Andrew Crooks[1,2,3], Peggy Agouris[1,4], Anthony Stefanidis[1,5]

1 Department of Geography and Geoinformation Science, George Mason University, Fairfax, Virginia, United States of America, 2 Center for Geoinformatics and Geospatial Intelligence, George Mason University, Fairfax, Virginia, United States of America, 3 Department of Computational and Data Sciences, George Mason University, Fairfax, Virginia, United States of America, 4 Center for Earth Observing and Space Research, College of Science, George Mason University, Fairfax, Virginia, United States of America, 5 Criminal Investigations and Network Analysis Center, George Mason University, Fairfax, Virginia, United States of America

* rmahabir@gmu.edu

Abstract

The last several decades have witnessed a shift in the way in which news is delivered and consumed by users. With the growth and advancements in mobile technologies, the Internet, and Web 2.0 technologies users are not only consumers of news, but also producers of online content. This has resulted in a novel and highly participatory cyber-physical news awareness ecosystem that fosters digital activism, in which volunteers contribute content to online communities. While studies have examined the various components of this news awareness ecosystem, little is still known about how news media coverage (and in particular digital media) impacts digital activism. In order to address this challenge and develop a greater understanding of it, this paper focuses on a specific form of digital activism, that of the production of digital geographical content through crowdsourcing efforts. Using refugee camps from around the world as a case study, we examine the relationship between news coverage (via Google news), search trends (via Google trends) and user edit contribution patterns in OpenStreetMap, a prominent geospatial data crowdsourcing platform. In addition, we compare and contrast these patterns with user edit patterns in Wikipedia, a well-known non-geospatial crowdsourcing platform. Using Google news and Google trends to derive a measure of thematic public awareness, our findings indicate that digital activism bursts tend to take place during periods of sustained build-up of public awareness deficit or surplus. These findings are in line with two prominent mass communication theories: agenda setting and corrective action, and suggest the emergence of a novel stimulus-awareness-activism framework in today's participatory digital age. Moreover, these findings further complement existing research examining the motivational factors that drive users to contribute to online collaborative communities. This paper brings us one step closer to

understanding the underlying mechanisms that drive digital activism in particular in the geospatial domain.

Introduction

Recent studies (e.g. [1]) indicate that the general public is increasingly using a combination of online platforms (e.g. news websites/apps and social media) and traditional news avenues (e.g. local and cable TV, radio, and print newspapers) to get its news. This results in a novel news ecosystem, supporting highly and actively engaged news consumers [2–4]. Within this news ecosystem, cyberspace is complementing traditional sources to communicate news [5–7], raise public awareness [8,9], and form and shape opinions and agendas [10–12] across a diverse range of issues.

This highly participatory cyber-physical news awareness ecosystem is fostering digital activism. Its manifestations range from encouraging and enabling the crowd to generate, publish, and disseminate focused information, to fostering the formation of online communities that jointly pursue activities in the cyberspace, and often the physical space too [13]. We are repeatedly witnessing this in the use of social media to support political activism in the streets of cities across the globe [14–19], the establishment of peer-to-peer information exchange communities [20,21], or even the formation of online communities to support cyberspace projects like Wikipedia [22]. While this powerful cycle linking information consumption with the raising of awareness and the resulting digital activism will continue to be essential to the future of our society, the mechanisms that drive this process still remain fuzzy.

One popular expression of such digital activism is in the form of volunteered contributions to online collaborative activities. In order to better understand the motivating factors that drive such activities, certain studies have focused on individual platforms separately. Nov [23] built on earlier work of Clary [24] that had identified the major motivational functions that drive volunteerism overall (e.g. altruistic values, social interaction, enhancing one's understanding and career-related benefits) to show that these general motivations apply to Wikipedia contributions as well. Xu and Li [22] classified these functions under two broad categories, namely the will to contribute content, and the interest in participating in that online community. Budhathoki and Haythornthwaite [25] examined the motivational factors that drive contributions on OpenStreetMap (hereafter referred to as OSM in this paper), a freely accessible and editable map of the world [26–28]. That study reported important positive factors relating to personal, yet the shared need to contribute to open source projects, the co-integration of individuals into open-source and geographic knowledge communities, along with the need to be attentive to participation taking places within the OSM community. It is also important to note that the motivation for contributors within these different communities can vary [29,30].

The above studies have been crucial in providing a better understanding of the various factors that drive people to become digitally active in online collaborative communities. However, relatively few studies have examined digital activism in the context of the previously described broader news awareness ecosystem. Zastro [31], for example, noted that during the Haitian earthquake in 2010, volunteer mappers used a combination of different news media (i.e. news reports, social media and text messages from survivors) to collect information on the damaged status of buildings. This information was then added to the OSM platform. Westrope et al. [32] examined the possible impact of media coverage on mapping activities following Typhoon Haiyan in the Philippines in 2013. That study showed that in areas that received substantial

media coverage (e.g. Tacloban City) 92% of the surveyed buildings had over-represented damage reported in the OSM edits. In contrast, in areas that received less media coverage only 56% to 76% of the buildings had over represented damage reported in the OSM edits. As the authors of that study suggest, the large over-representation in damage buildings in Tacloban City could be attributed, at least in part, to the priority given to this city by the Humanitarian OpenStreetMap Team (HOT) OSM community. In a related study, Dittus et al., [33], compared OSM activity for various Humanitarian HOT events, showing that the two most publicized media events, Typhoon Haiyan, and the Nepal earthquake in 2015, had the largest newcomer mapping recruitment rates. The authors of that study suggested that news coverage of these events may have assisted in attracting a much larger number of newcomers. Similar findings of the influence of media coverage and an increase in newcomer enrollment in the OSM platform were reported by Begin et al., [34]. While these studies shed some light on the possible relation between OSM activity and news media coverage, a more in-depth analysis of this relation is still needed.

The pursuit of a deeper understanding of the relationship between the different components of the complex news awareness ecosystem and digital activism is becoming a substantial research challenge. Wikipedia edit patterns, for example, have been compared to breaking news [35], and news media coverage of an issue to increases in searches for and edits of related articles in Wikipedia [36]. Such patterns have also been used to define entity-specific news tickers and timelines [37], with page views further used to detect popular topics related to users' interest [38]. Google Trends [39], a tool for measuring public agenda [40,41], has also been found to be useful for quantifying trading behavior in financial markets [42,43], for disease surveillance and health care research [44,45], and as a tool for behavior analytics, such as predicting non-cigarette tobacco use [46]. Moreover, Ratkiewicz et al., [47] show that bursts of online search volume activity extracted from Google Trends tend to be correlated to similar bursts of editing activity on Wikipedia. Such studies highlight just how complex this news awareness ecosystem is.

In the context of this paper we use the term news awareness ecosystem to refer to the ensemble of sources, ranging from traditional newsrooms to grassroots citizen journalism, that provides the public with possible exposure to news, and from which public awareness can emerge. A key characteristic of this ecosystem is that rather than being monolithic it is multifaceted, multimodal, and evolving. This term builds on the idea of the "news ecosystem" that has become prevalent in recent years. While an agreed upon definition of this term is still lacking, several attempts towards the construction of such a definition have been made. For instance, Anderson [48] provides a brief genealogy of the term and its origins, and discusses several concepts that may support the construction of such a definition. Alternatively, Picard [49] provides a detailed discussion of the forces that reshaped news journalism into what is referred to as the new (news) ecosystem.

Taking a more non-contemporaneous view of the news cycle, Nghiem et al., [50] showed that, depending on the specific topic, news coverage can be responsive to or lag an occurring trend (as defined by Google Trends search volume data). In some cases, albeit very few, a build-up of news activity was also followed by a decrease in search volume. Such responsive behavior of news patterns has been documented for trending topics in social media as well [51,52]. Althoff et al., [53] compared trending data from Google (Search, News and Trends), Twitter and Wikipedia to show that combined, these different data platforms can be used to forecast trending topics, even though temporal activity patterns (comprising spikes and cumulative build-up) tend to vary across these platforms. Other work by Al Emadi et al., [54] showed the strong response of the online mapping communities in the aftermath of natural

disasters, e.g. in the form of increased online volunteering activity in the MicroMapper [55] crowdsourcing platform.

Digital activism and geography have long been intertwined, giving rise to a range of online user activities, which in the context of this paper we refer to as geo-activism (in the remainder of the paper we will use the term digital- and geo-activism interchangeably). From crowd-sourcing crisis information mapping, to support disaster relief and recovery efforts [56,57] and collectively editing geographical data in OSM [27], volunteered geographical information (VGI) has emerged as a substantial new mechanism for the general public to contribute geographical content by mapping roads, buildings, and other artifacts [58–60]. At the same time, it is important to recognize that the general public is also increasingly engaged in geo-activism while being exposed to the news media. As suggested in previous studies [61–63], many forms of online news media have the power to shape public opinion, affect other media sources, and foster engagement.

Consequently, a key premise of our work is that geo-activism should be considered in the context of the news media rather than as a separate, independent phenomenon, in order to advance our understanding of the mechanisms that drive it. This would allow us, for example, to explore how the ebb and flow of information in the public media sphere affects digital geo-activism. At a time where VGI is establishing itself as a rich supplementary—and sometimes only—source of geospatial information [64, 65], a better understanding of the mechanisms driving geo-activism will allow us to further harness its power.

Accordingly, our objective in this paper is to study links between media coverage and geo-activism by focusing on the question of *geographical saliency* within the news awareness cycle. The basic argument is that news features of broad community interest but with a certain, narrow geographical footprint build up over time awareness and interest for that location, leading to bursts of volunteered geographical contributions for it. We pursue this task by comparing patterns of online volunteered geographical edits in OSM to relevant news stories, and to corresponding edit patterns in Wikipedia. Therefore, the research question at the core of this paper is whether extreme instances of public awareness deficit or surplus (resulting from sustained relevant trends) tend to trigger digital geo-activism bursts.

Towards this goal, we present in this paper a case study that focuses on several refugee camps around the world to determine whether salient features in news media affect digital activism in OSM. This line of inquiry allows us to examine the possible association and interconnectedness between macro level global awareness and news coverage and digital activism, with a hyperlocal geographical focus.

A number of factors render refugee camps particularly suitable for such a study. First, they match very well the above-mentioned motivational factors that drive digital activism as they were identified by Nov [23]. Second, compared to previously studied events like natural disasters, refugee camps differ in the sense that they are not abrupt events that occur over a period of few hours (e.g. earthquake) or days (e.g. flooding), which would render the relationship between media coverage and digital activism trivial, as event, news coverage, and response practically coincide. Instead, these refugee camps are set up and operate over a period of years [60] and public awareness to them builds over time. Third, they offer the advantage of having a geographically distinct footprint in terms of size and location relevant to their surroundings while being associated with a narrow thematic focus at the same time (unlike large cities, which may have numerous themes occurring at the same time). Finally, in the aftermath of massive displacement of populations due to civil war or conflict (e.g. the Syrian crisis that began in 2011 and other on-going ethnic tensions), refugee camps represent a topic that is becoming increasingly important in terms of media coverage and public awareness [66].

The remainder of this paper is organized as follows. In Section 2, the datasets and methods used to address our research question are presented followed by an in-depth discussion of the results in Section 3. Finally, Section 4 provides a discussion of our results and concludes with an outlook of future work.

Materials and methods

Refugee camps

The current concentration of refugees is highest in African and Middle Eastern countries compared to other world regions [67]. These regions also contain the most populous camps as well [68]. Accordingly, a set of 8 refugee camps was selected for our study, located in Africa, Middle East, and Europe, as shown in Fig 1. In addition to reflecting the global distribution of refugees, these camps offer diversity in terms of their size, population, and date of establishment. Camp age (date of establishment) is of particular interest, as it allows us to study camps at different stages of news cycles and interest.

Table 1 shows more detailed information for each camp including their established date and camp status, their size (area), population, and population density (See S1 File for references). As expected, larger camps tend to host more refugees, and the correlation between camp size and population has a Pearson r value of 0.70 with a p-value of 0.03 (with Kakuma being the one exception). However, in general, there is no clear observable association between the age of the camps and their population density or size.

Data sources

To analyze patterns of geo-activism as it relates to these refugee camps, we used the patterns of relevant information contributions in OSM. OSM is a prototypical example of a VGI platform allowing anyone to map features on the Earth's surface. It was launched in July of 2004 and presently, as of August 2018, has almost 5 million registered users [69]. While being general in its scope, OSM has grown over time to be a particularly rich source of geographical data, and especially so in support of humanitarian response activities. Some recent examples include OSM contributions relevant to the Ebola outbreak in West Africa in 2014 [70], the Nepal earthquake in 2015 [71], and slums in Sub-Saharan Africa [65].

As OSM is a collaborative platform of information contribution, it tends to exhibit activity patterns similar to those found in other non-geographical platforms. For example, a common thread in many online collaborative communities is that the level of participation is not even: few users tend to contribute massive amounts of information, whereas the large majority tends to contribute less. Previous studies have confirmed that this pattern is also present in the OSM community [72,73]. Similar patterns of contributor activity have also been found with other online crowdsourcing communities such as Wikipedia [74–77], which are often characterized by bursts of activity [78,79]. When it comes to OSM, the spatial and temporal frequency of contributions can vary based on numerous factors, including, the diversity of contributors [27,80], issues associated with the digital divide [81–83], the social structure of contributor communities [84–85], the direct intervention of social groups such as mapping parties [86,87], and bulk imports into the OSM platform [88], among others.

For the purpose of our study, we used OSM data that were extracted from the *planet history file* (https://planet.openstreetmap.org/planet/full-history), which contains all OSM edits for the entire world. An OSM edit herein is defined as any create, delete or modify operation to any OSM node, way or relation feature. The OSM planet history file also contains other useful information, such as the location (latitude and longitude) of each node, object versions (from October 2007 onwards), contributor user ID, and timestamp. Refugee camp OSM data were
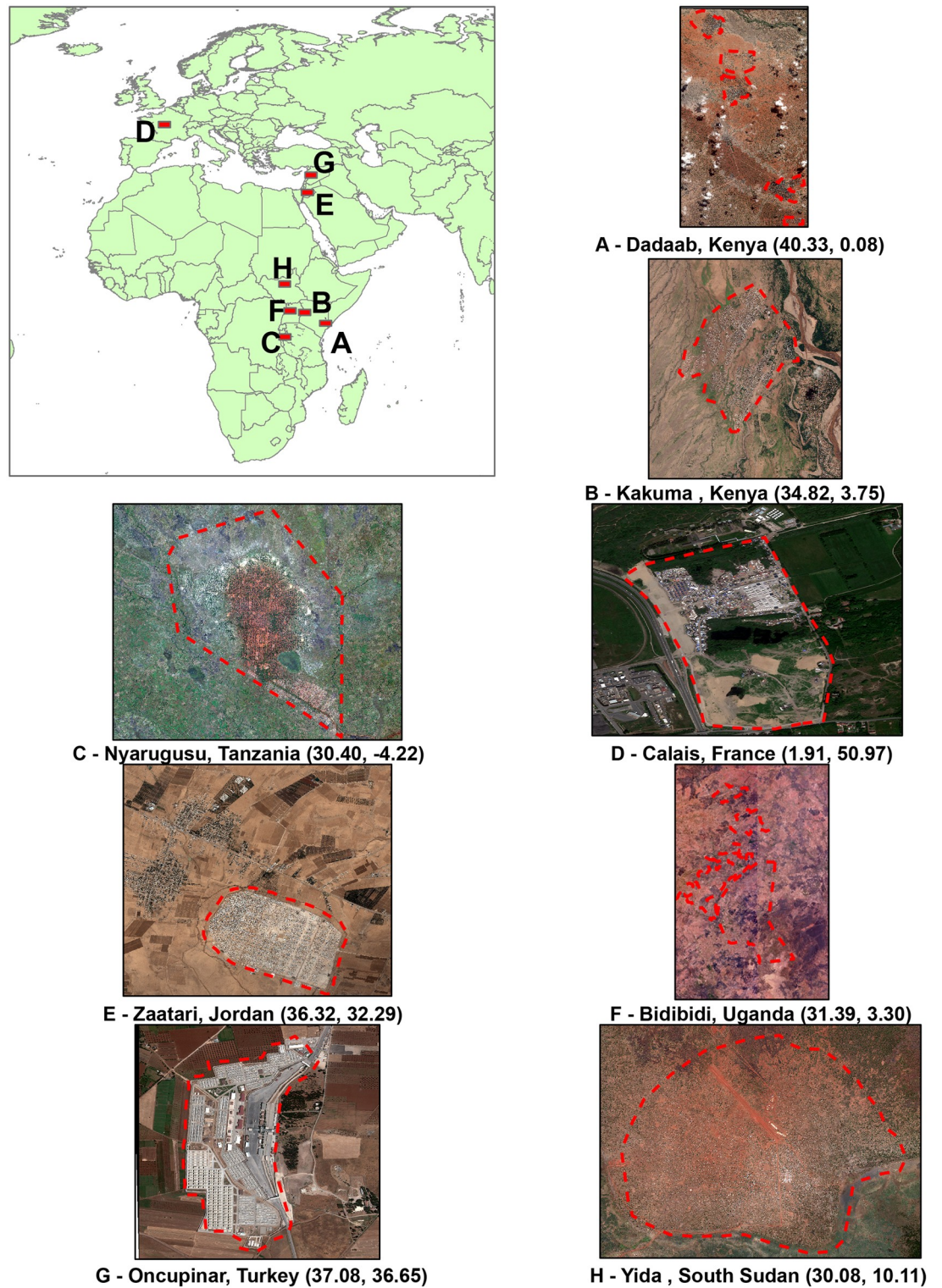
Fig 1. Study areas (centroid location of camp). Satellite image courtesy of the DigitalGlobe Foundation.

Table 1. Refugee camps.

Camp	Country	Established Date/Status	Area (km^2)	Population	Population Density (People per km^2)
Dadaab	Kenya	1992 with an addition in 2011 / Active	40.7	242,998	5,970
Kakuma	Kenya	1991/Active	6.4	171,085	26,732
Nyarugusu	Tanzania	1996/Active	25.9	78,519	3,032
Calais	France	2015/Closed on October 2016	0.61	9,000	14,754
Yida	South Sudan	2011/Active	20.4	55,012	2,697
Bidibidi	Uganda	2016/Active	360.9	270,000	748
Oncupinar	Turkey	2012/Active	0.71	15,000	21,127
Zaatari	Jordan	2012/Active	6.1	79,827	13,086

clipped from the history file using spatial and temporal parameters. In order to delineate the footprint of each camp we use polygons that demarcated their spatial extent, and use these polygons to select the corresponding edits. Camp polygons were manually traced using high resolution satellite imagery over each camp. In terms of time window, we selected edits during the current decade (01/01/2010 to 05/31/2017). From among all our camps, only Bidibidi had edits (only 2) prior to our study window, and so in practice our study addresses the whole OSM history of these camps.

In conjunction with the collection of OSM edit activity, edit activity was collected for the Wikipedia page entries that correspond to each of the refugee camps in Table 1. A Wikipedia edit in the context of this research refers to any change (i.e. creation, modification or deletion operation) made to the content of a Wikipedia page. In our study the edit activity of each page was collected from the history page of each camp's Wikipedia page using the same time window parameters noted above for OSM. Only English pages were used in this study, a caveat that will we will later revisit in the Discussion section. Wikipedia is the world's largest free online encyclopedia, allowing anyone to create new and edit existing articles. As of August 2018, more than 34 million users were registered with almost 46 million pages, and with the number of entries now approaching 6 million [89]. A key motivation to explore the edit activity of Wikipedia alongside the edit activity of OSM is that while both platforms rely on digital activism through crowdsourcing efforts, the former focuses on digital activism in a more general sense while the latter focuses specifically on geo-activism. This difference will enable our analysis to compare and contrast digital activism patterns in two substantially different crowdsourcing environments.

In order to analyze relevant news media activity to OSM edit activities we also collected relevant data from Google News and Google Trends. While Google News conveys expressions of media coverage, Google Trends is considered here as a proxy for public interest. News coverage data were captured from Google News [90], an online news aggregator for up-to-date news stories from all over the world. As of 2012, Google stated that it draws from more than 50,000 news sources with more than one billion unique users connected each week to its news content [91]. Prior research has also compared news extracted from Google with other platforms such as LexisNexis, showing that Google News provided broader worldwide coverage than its counterparts [92]. Google News articles were searched using the keywords "*name of camp*" AND "*refugee*" to help filter non-relevant stories. Articles were further filtered manually by the authors as needed, to remove irrelevant data. Between 6% and 26% (average of 13%) of articles were removed during the data cleaning process. Filtering was initially done by one of the authors who collected the data, and later on two other authors assisted to further remove articles that were deemed non-relevant. The remaining relevant news articles were then binned into daily counts.

To capture information on public interest, we used Google Trends data. Google Trends provides information on how often a particular keyword was searched using Google's web search engine. This data has also been used in previous studies as a proxy measure for public interest in a variety of topics [40,50,93,94]. In order to access Google Trends data for each camp, we used the corresponding online portal and the camp's name as the keyword, to capture searched volumes for each camp. This data provided by Google, is available as a normalized series of values between 0 and 100, based on the popularity of users' search interest within the specified temporal range. A summary of the various steps used to extract and preprocess the different data sources (OSM, Wikipedia, Google News, and Google Trends) are presented in Fig 2. Further, in S2 File, we show the total number of OSM edits, Wikipedia edits and Google News articles at the daily level for the period 01/01/2010 to 05/31/2017. With respect to Google Trends, data at the daily level is limited to a temporal search window within 269 days. Following this, the data is provided at weekly and then the monthly aggregated level depending on the temporal search window. As such, the Google Trends data is not included in S2 File. An initial visual comparison of the data at the daily level did not reveal any apparent patterns across the different data sources.

Overview of multiscale and multiplatform comparison approach

Given the objective of this research, our approach is based on using four basic measures to explore the possible associations between news media coverage and crowdsourced activity: Google Trends index, Google News volume, OSM edits activity, and Wikipedia edits activity. While the first two measures serve as a proxy of the general public's interest and availability of news around specific themes, the latter two measures serve as a proxy of the level of digital activism around specific thematic features over space and time. A key premise of our approach is that comparing and contrasting these measures can provide additional insight into the possible relationship between the ebb and flow of news and activity patterns in digital activism in platforms such as OSM and Wikipedia.

While gauging the levels of digital activism is relatively straightforward (e.g. tracking the number of edits in Wikipedia), measuring the ebb and flow of news in the media ecosystem is more challenging as there is no such single readily available measure. To overcome this issue, we propose to consider the difference between the normalized Google Trends index and the normalized Google News metric as a measure of the overall *public awareness* with respect to a specific theme (or a set of themes). Here, we refer to the term awareness simply as the "knowledge that something exists, or understanding of a situation or subject at the present time based on information or experience" [95]. The reasoning behind using this difference measure is that while Google News represents the availability of news (i.e. information) regarding a theme, Google trends represents the public's active pursuit of information about the theme. As news stories emerge, evolve, re-emerge, and eventually subside in the media, the public's information seeking activities may increase, decrease, or remain unchanged over time [50]. Consequently, a surplus (or deficit) of *public awareness* can be built-up over time through the availability (or unavailability) of news stories that consistently increase (or reduce) the public's online information seeking activities. A prolonged period of awareness surplus growth would be one where the normalized metrics of Google Trends grow faster than the corresponding metrics of Google News, implying that public interest on the topic grows faster than news coverage. Conversely, a prolonged period of awareness deficit would be one where news coverage outpaces public search interest. In this context of this paper, a min-max approach was used to rescale all four data metrics to values between 0 and 1.
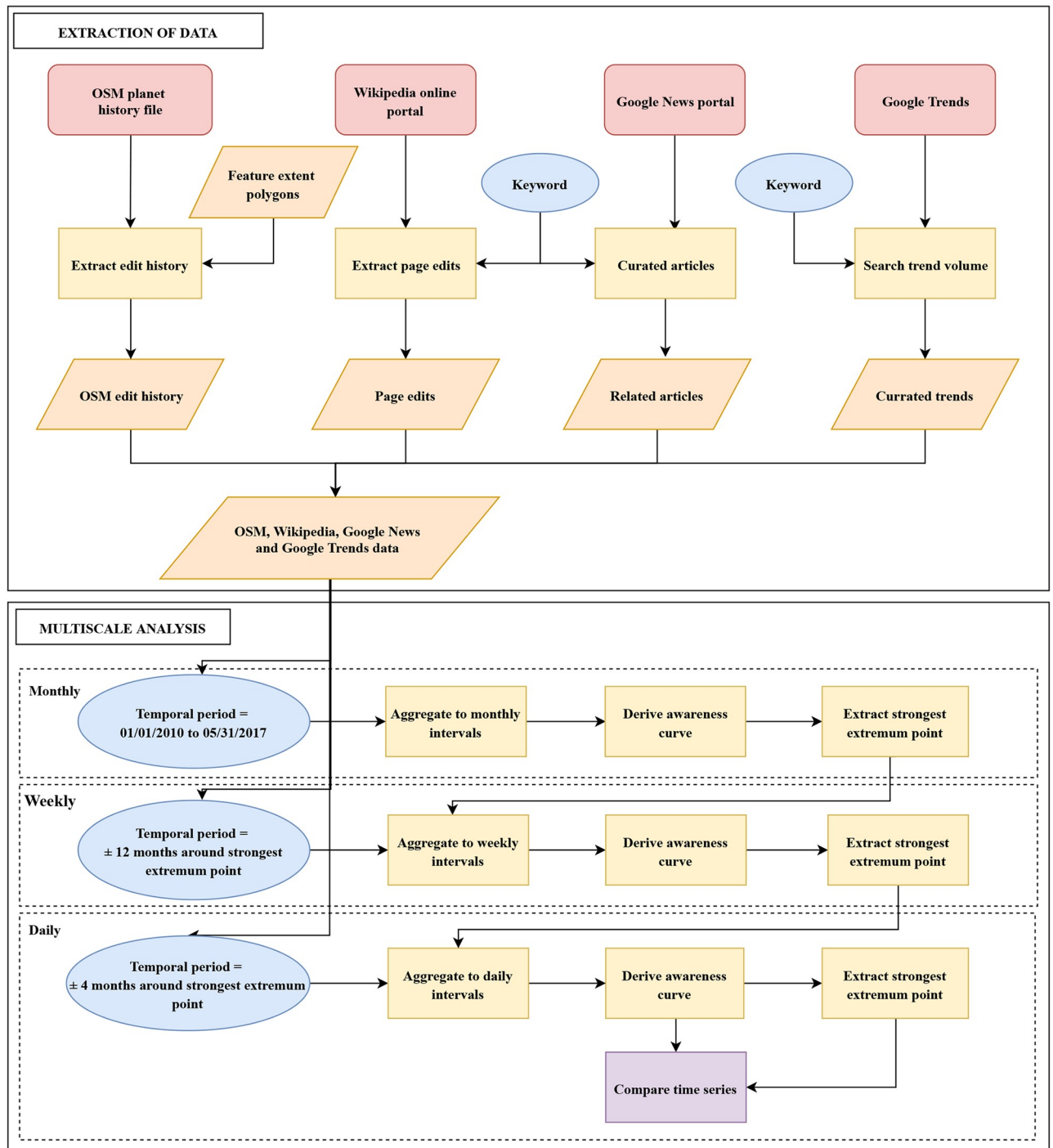
Fig 2. Data extraction and preprocessing. In the data extraction step, data on OSM, Wikipedia, Google News, and Google Trends is collected from their various online platforms. The multiscale analysis step involves a progressive refinement approach: Data is first examined at the monthly level and the strongest extremum points on the awareness curve are identified. These points are then assigned a score that is comprised of its magnitude and a weight that is proportional to the duration of consistent trend along the awareness curve prior to the detected extremum point. A search time window is then defined on the awareness peak with the highest score and used for working with data at the weekly analysis level. A similar approach is used to identify a search window for working with data at the next (finer) temporal granular analysis level.

The ability to track trends in public awareness overall enables us to explore the potential associations between such trends—and in particular trend changes—in awareness and digital activism. Specifically, our interest lies in the question of *how are trend changes in public awareness related to user activity in OSM and Wikipedia?* The emerging argument from this question is that media coverage drives participation, by informing the general public of evolving/developing situations, thus setting an agenda and leading to digital activism. Since such a mechanism can occur at different time scales, our approach is based on consecutively examining the different measures at 3 levels of time granularity, namely monthly, weekly and daily, in order to identify possible associations at the finest temporal granularity considered here. Starting from the monthly level, extremum points in public awareness activity are detected, and the strongest extremum point is identified. The time stamp of this point is then used together with a search time window around it to define a new search interval at the next (weekly) granularity level. This process is then repeated for identifying a search time window at the daily time granularity, in which possible relations between public awareness trend changes and OSM and Wikipedia activity are explored. A more detailed description of the extremum points detection and selection process is provided below.

Our analysis process begins by examining the public awareness curve at the monthly level and then proceeds using a progressive refinement mechanism towards an examination of data at the daily level. In order to capture broad trends in public awareness level, we model the public awareness data using multivariate adaptive regression splines (MARS). MARS is particularly suitable for this type of data since it makes no assumptions about the underlying distribution of the data and provides an intuitive approach for understanding the intrinsic complicated data mapping in high-dimensional data patterns [96]. MARS is represented as a combination of basis functions expressed as [97]:

$$y = \beta_0 + \sum_{m=1}^{M} \beta_m h_m(X) \qquad (1)$$

where y is the dependent variable (i.e. our public awareness measure), X is the independent variable (time), β_0 is the intercept parameter, and β_m is the coefficient applied to each basis function $h_m(X)$, which are summed over M non-constant terms used for defining the number of basis functions. In this case, M is determined in a data-driven manner based on the optimum number of basis functions required to fit the data. While MARS supports polynomial basis functions, in our implementation linear basis functions are used for simplicity.

Using the derived MARS model, extremum points in the public awareness curve are detected by estimating the second derivative $f''(x)$ as a central finite difference approximation:

$$f''(x) \approx \frac{f(x_i + \Delta x) - 2f(x_i) + f(x_i - \Delta x)}{\Delta x^2} \qquad (2)$$

where x_i represents the ith value of x in the data series defined by the predicted values in Eq 1, Δx is the step size in temporal units (i.e. monthly, weekly or daily dependent on the scale of analysis) around point x_i, where $i = 0,1,2,\ldots,n$. At each temporal granularity level Δx is set to 1 time unit, e.g. at the monthly level a Δx value of 1 month is used.

As the detection of extremum points along the public awareness curve may result in multiple local maxima or minima points, a pruning process is applied in order to capture the most prominent extremum points. This process begins by identifying all extremum points and then for each point suppressing other weaker extremum points (in terms of their absolute magnitude) within a local window centered around each top extremum point. In our analysis this was set to ±1.5 months, ±1 week, and 0 for the monthly, weekly and daily granularity, respectively. In the case of Google Trends, this data was first captured at weekly granularity to

identify local windows around each extremum point at the monthly analysis level. Following this, at the next level of analysis (i.e. weekly), the Google Trends data was captured at daily granularity. No local windows were used at the daily analysis level.

Following the pruning process, each of the remaining extremum points (in terms of their absolute magnitude) is assigned a score that is comprised of its magnitude and a weight that is proportional to the duration of consistent trend (positive or negative) along the awareness curve prior to the detected extremum point. This weighting scheme is guided by previous work [98], which suggested that the longer a theme is circulated within the media, the greater its potential for influencing the public agenda (which may result in digital activism). Based on these calculated scores, the awareness peak with the highest score (strongest extremum point) is selected, and is then used for defining a search time interval ΔT at the next (finer) temporal granularity level.

The selection of ΔT in our analysis is guided by prior research on agenda setting theory in communication [99], in which the question of the time that it takes for the public to respond to news stories was explored. Such prior work suggested that it may take as little as a few days [100] to as much as 2 to 6 months [98,101] for changes in the media agenda to become fully realized into public agenda. In other studies, e.g. [93], a period of 50–70 days was suggested. Informed by these prior studies, our analysis is based on repeating the search for the for the highest scoring awareness peak using a progressively refined ΔT. Specifically, ΔT was set to ±12 months and ±4 months around each highest scoring awareness peak in the monthly and weekly time granularity, respectively.

As our objective relates to the comparison of the awareness curve to the two crowdsourcing platforms, it is also necessary to detect extremum points in the OSM and Wikipedia edit activity data. For convenience, and due to the overall stepwise nature of this data, we consider the cumulative edit activity time series for these two platforms and approximate each edit activity curve using the Douglas-Peucker line simplification algorithm [102]. Then, using the simplified edit activity curves, we detect the top 5 extremum points ranked by their overall magnitude. In some cases, fewer than 5 significant peaks were extracted as a result of the line simplification process.

Results

Given the objectives of this paper and the proposed analysis approach, this section summarizes the results that were derived along three main themes. First, we outline the key trends in the edit activity for the eight camp sites. Following this, we present the results of an analysis of the edit activity patterns in the local context of each camp site. Finally, we examine how trends in public awareness relate to OSM and Wikipedia edit activity patterns.

Edit activity trends in OSM

The overall number of OSM edits for all 8 refugee camps for the period 01/01/2010 to 05/31/2017 is shown in Table 2. Comparing these numbers with the data in Table 1 reveals a strong correlation between camp size (spatial extent) and number of OSM edits, with a Pearson r correlation of 0.89 and a p-value of 0.002.

In order to delve into the driving forces behind public participation in OSM we further look at the temporal variations of these contributions. Fig 3 shows the patterns of temporal activity in OSM edits for the 8 camps selected. The data in Fig 3 has been normalized to values between 0 and 1 to account for different scales of edit activity between camps.

Fig 3 shows that edit activity in OSM tend to occur in bursts (spikes in the charts), similar to the what was observed in Wikipedia [47]. This observation is consistent with Barabasi [103],

Table 2. Total OSM edits per camp for the study period (01/01/2010–05/31/2017).

Camps	Total OSM Edits
Dadaab	31,283
Kakuma	11,147
Nyarugusu	74,416
Calais	5,965
Yida	84,053
Bidibidi	198,916
Oncupinar	708
Zaatari	69,981

who suggested that such an activity pattern tend to follow a non-Poisson pattern, which is common in human activities in general. Following this theory, edit activity bursts could then be seen as following a natural pattern of contributors' interest (as a measure of priority) in mapping refugee camps. Furthermore, as such OSM edit activity bursts represent epochs in time in which users' activity was focused on specific localized camp sites, one could argue that these bursts represent instances when these sites become salient features/attention landmarks [100] in geographic space. A more detailed discussion of this saliency property is provided in the next section.

In addition, it is interesting to observe that the OSM edit activity bursts in Fig 3 do not tend to coincide in time across camps. This is expected as it is rather unlikely that users that edit different camp sites will do so at the exact same time, and suggests that geo-activism in OSM around this theme tends to be asynchronous across different locations. This asynchronous nature could also be attributed, in part, to the specific history of each camp site. Some camps, such as Bidibidi, exhibit OSM edit activity bursts towards the end of our study period, while more established camps, such as Dadaab, exhibit several bursts of activity spread throughout the study period. For camps such as Nyarugusu, a single distinct edit activity burst is apparent. Finally, there were no clearly observed co-occurrences of OSM edit activity in Fig 3 and camp establishment dates as noted in Table 1.

Geographical saliency: Camps as local activity hotspots

In order to assess the degree to which a refugee camp becomes a local geographical salient artifact, overshadowing interest on its immediate surroundings, we evaluate the extent to which OSM editing activity within their boundaries exceeds the editorial activity in their immediate surroundings. In that sense, camps then become local activity "hotspots" of digital geo-activism. In order to estimate the OSM editing activity both within and around the camps, we define a set of 4 zones for each camp: a camp zone Z_0 and 3 surrounding aerial zones, Z_1, Z_2, and Z_3. The relationship between camp and surrounding zones is shown Fig 4.

Formally, zone Z_0 is defined as the spatial footprint of the camp, as captured by the set of (one or more) polygons P = {p_1, p_2, . . ., p_n} (n≥1) that delineate it. Based on this, zone Z_0 is defined as:

$$Z_0 = \cup \alpha(p_i), i = 1, 2, \ldots, n \qquad (3)$$

where $\alpha(p_i)$ is the area enclosed by polygon p_i, and $\cup$ is the union operator. Given P, the set of minimum bounding boxes of p_1, p_2, . . ., p_n are then derived, and the maximum side length (width or height) d_{max} among all bounding boxes is found. In addition, the minimum bounding box B that encloses P is also derived. Using d_{max} and B the zones Z_1, Z_2, and Z_3
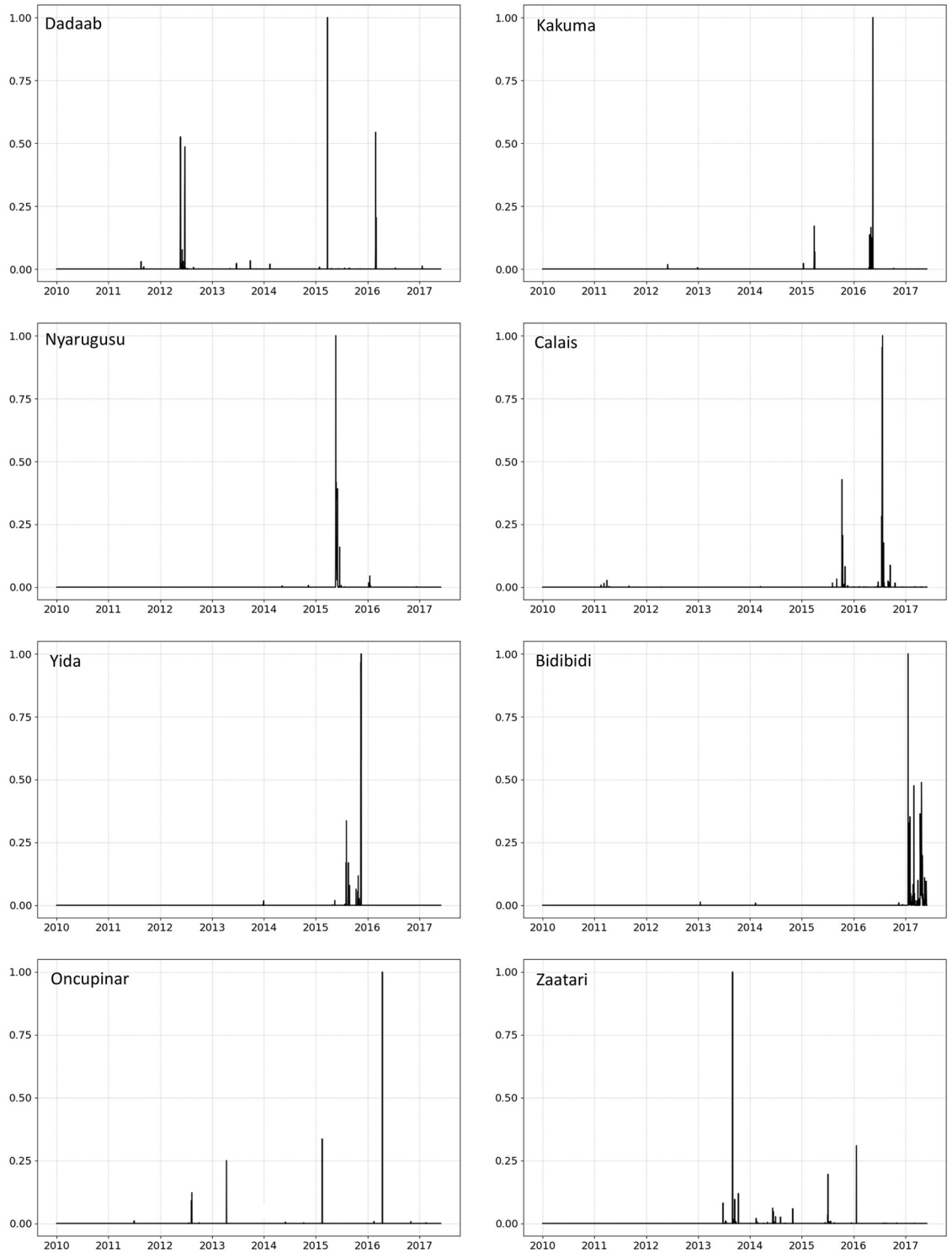
Fig 3. OSM edit patterns for the 8 camps studied (01/01/2010–05/31/2017).

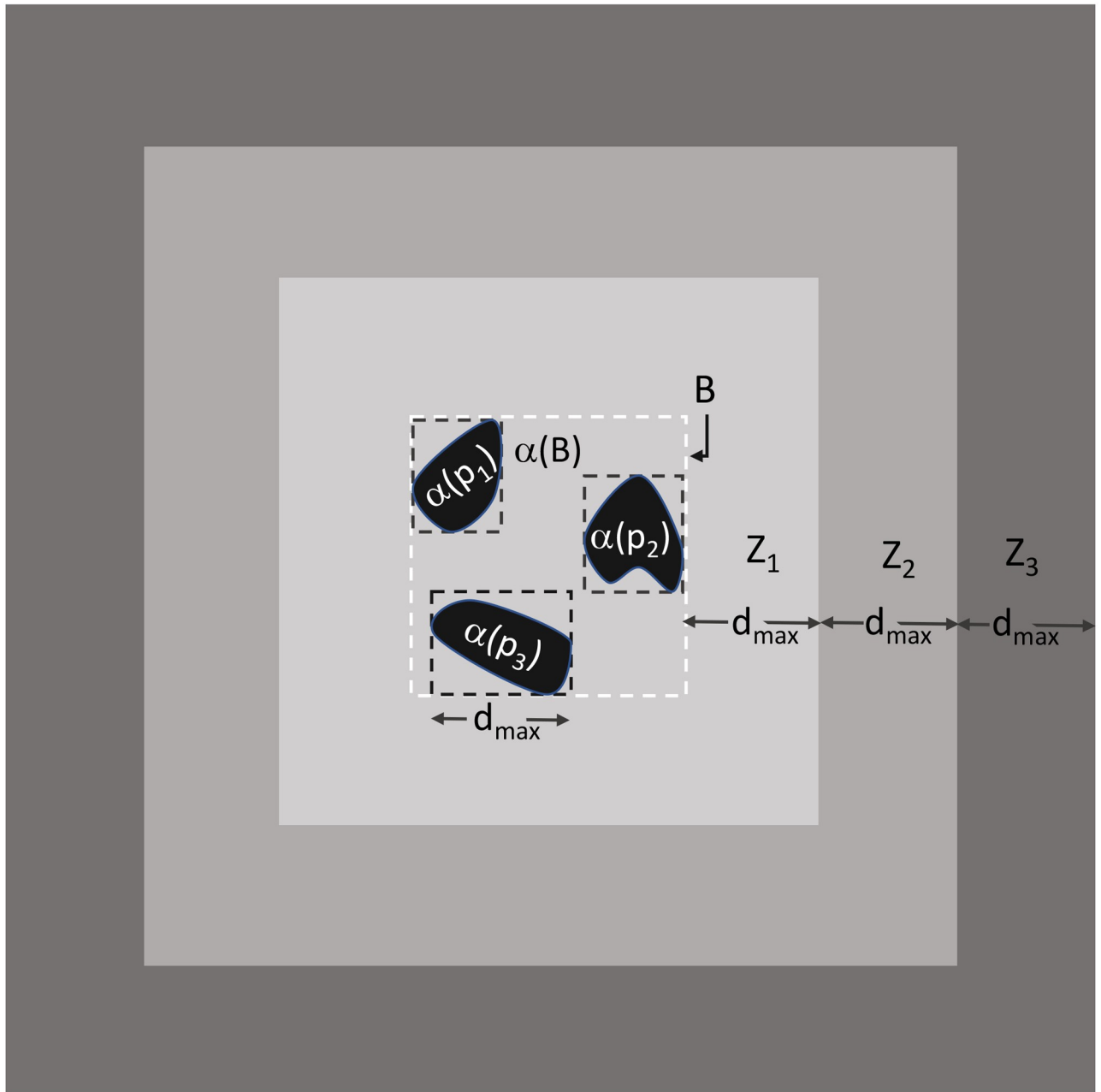https://doi.org/10.1371/journal.pone.0206825.g003

Fig 4. A schematic example of a camp (aggregate of black polygons), its minimum bounding box (dashed white line B) and its surrounding zones Z_1 (light gray), Z_2 (aggregate of light and mid gray), and Z_3 (aggregate of Z_2 and darkest gray).

(respectively) are derived by repeatedly applying a dilation operator ($\oplus$) on the area enclosed by B (denoted by $\alpha(B)$) using a square structuring element D with a dimension d_{max} as follows:

$$Z_1 = \alpha(B) \oplus D + \alpha(B) - Z_0; \;\; Z_2 = Z_1 \oplus D - Z_0; \;\; Z_3 = Z_2 \oplus D - Z_0 \tag{4}$$

We use these zones to compare OSM editing activity within the camp versus its 3 surrounding zones (Z_1, Z_2, Z_3) during a ±4 month period around the strongest extremum point extracted from the weekly data, and examined at the daily granularity for each camp. The

Table 3. Temporal window (±4 month period) around the strongest extremum point extracted from the weekly data at the daily granularity.

Site	Start date	Stop date
Dadaab	10/16/14	6/15/15
Kakuma	10/6/16	6/5/17
Nyarugusu	1/25/15	9/24/15
Calais	5/7/15	1/6/16
Yida	9/25/15	5/24/16
Bidibidi	9/23/16	5/22/17
Oncupinar	1/30/16	9/29/16
Zaatari	5/19/16	1/18/17

https://doi.org/10.1371/journal.pone.0206825.t003

temporal window used for each refugee camp is shown in Table 3 with the results from this analysis summarized in S3 File. These results exhibit a sharp decline in editing activity along the transition from the camps outwards, to their surrounding zones. On average, for five of the eight camps the drop in number of OSM edits from Zone Z_0 to Z_1 was 93%. The remaining three camps, namely Kakuma, Calais and Bidibidi, have edits in Z_1 exceeding edits within Z_0. In the case of Calais, this finding can be explained by the camp's neighboring synonymous city being part of the Calais Z_1 zone. Similarly, in the case of Bidibidi, the increase in the number of OSM edits can be attributed to the fact that Bidibidi is surrounded by other refugee camps that are characterized by substantial edit activity, and are included in Bidibidi's corresponding Zones Z_1 to Z_3.

The drop in the number of OSM edits becomes even more pronounced when considering the normalized edit metrics per area for each zone. On average, for the eight camps studied here, Z_1, Z_2, and Z_3 zones cover an area that is 18, 54, and 108 times larger than the area of Z_0, respectively. In order to account for these zone area variations, the number of OSM edits is normalized to be per km^2, as shown in S3 File. With the exception of Kakuma, the data shows that on average the number of OSM edits per km^2 drops as one moves from Z_0 outwards by 96.37% (Z_1), 97.96% (Z_2), and 98.42% (Z_3). This pattern was also similar for Kakuma, however, the drop in the number of edits per km^2 was much lower, moving from Z_0 outwards by 2.71% (Z_1), 63.99% (Z_2), and 81.91% (Z_3). These results, therefore, support the notion that the camps are indeed serving as local activity hotspots, attracting OSM edits from the corresponding volunteer community beyond what would be expected by their surrounding areas.

Similar to the number of OSM edits, an analysis of the number of contributors per km^2 was carried out, as summarized in S3 File. These results exhibit a trend similar to the one found in the number of OSM edits per km^2: the number of contributors per km^2 also drops as one moves from Z_0 outwards by 95.07% (Z_1), 97.59% (Z_2), and 98.09% (Z_3). Interestingly, in the case of Calais, the number of edits per km^2 is increasing moving from Z_0 outwards. Once again, this reversal in trend can be explained by the proximity of this refugee camp to a large urban area in a developed country.

In order to complete the saliency analysis, S3 File also lists the number of individual camp contributors who were active in the surrounding zones. As can be seen from these results, overall, the number of contributors who were active in the each of the camp zones (zone Z_0) who were also active in other zones is approximately 40% (the values range between 25%-50%), further suggesting the role of camps as local OSM edit activity hotspots.

Public awareness versus OSM and Wikipedia edit activity

In order to assess the possible relationship between news media coverage and digital activism in OSM, we compare OSM edit activity to three additional data sources, namely Wikipedia edits, Google News and Google Trends for each camp using the progressive refinement approach described earlier. As stated previously, in the context of this comparison, Google News is used as an indicator of media coverage, conveying how frequently a refugee camp appeared in news. In contrast, Google Trends is used as an indicator of broad public interest in this camp. Finally, Wikipedia edit activity represents an example of a non-geographic form of digital activism.

Fig 5 shows the time series of all four data sources (OSM edits, Wikipedia edits, Google News items, and Google Trends indicator) for each of the eight camps during a ±4 months period around the strongest extremum point of each camp (at the weekly temporal granularity–Table 3). Based on these time series data, the public awareness curve was calculated (as the cumulative difference of Google Trends and Google News activity) for each camp and the strongest extremum points were detected. Then, using the progressive refinement approach presented earlier, we examined the relationship between the public awareness curve and the OSM and Wikipedia edit activity in a time window of ±4 months around the extremum point. Fig 6 depicts the public awareness curve (magenta line) along with the cumulative OSM and Wikipedia edit activity (black and red lines, respectively). In these graphs we also show splines (dashed blue lines) fitted to the public awareness curves to better visualize the overall trends in these curves. In the context of our approach, because we analyze extreme cases of public awareness, any significant deviation between the two input variables (i.e. Google Trends and Google News) at the daily analysis level is viewed as a possible predictor of activism activity.

As the graphs in Fig 6 show, digital activism bursts tend to be associated with periods of consistent build-up of surplus or deficit in public awareness. This tendency is particularly evident in the Dadaab, Nyarugusu and Yida camps, where OSM edit activity bursts tend to overlap with public awareness surplus, and in the Bidibidi camp, where OSM edit activity bursts coincide with public awareness deficit. Such patterns of activity are also evident with respect to Wikipedia edits for some camps: for example, Dadaab, Nyarugusu and Bidibidi. However, in other camps, namely Kakuma, Calais and Zaatari, Wikipedia edits exhibit a rather prolonged sustained effort of edit activity compared to the activity bursts in OSM (note that for Yida and Oncupinar no Wikipedia edits were made during the ±4 months period around the strongest extremum point).

To further examine the association between public awareness trends and OSM and Wikipedia edit activity, we derived the time gap between the most significant extremum point in OSM and Wikipedia extremum points to the closest extremum point in the public awareness curve of each camp at the three temporal granularity levels. The time gaps that were found at the monthly, weekly, and daily time granularities as a result of the progressive refinement process along with the range of time gap values across the eight camp sites are provided in Table 4. As can be seen from this table, at the finest (daily) temporal granularity, the average time gap is approximately between 11 and 12 days.

Discussion

Today's age of the participatory news consumer [104] has been steadily blurring the lines between digital content consumption and production [105,106]. An emerging manifestation of this change is the bridging of the gap between the omnipresence of news in one's daily life and one's resulting expression of activism [107,108]. Such expressions of activism have previously been studied in the context of the shaping of public policy (e.g. [109,110]), political
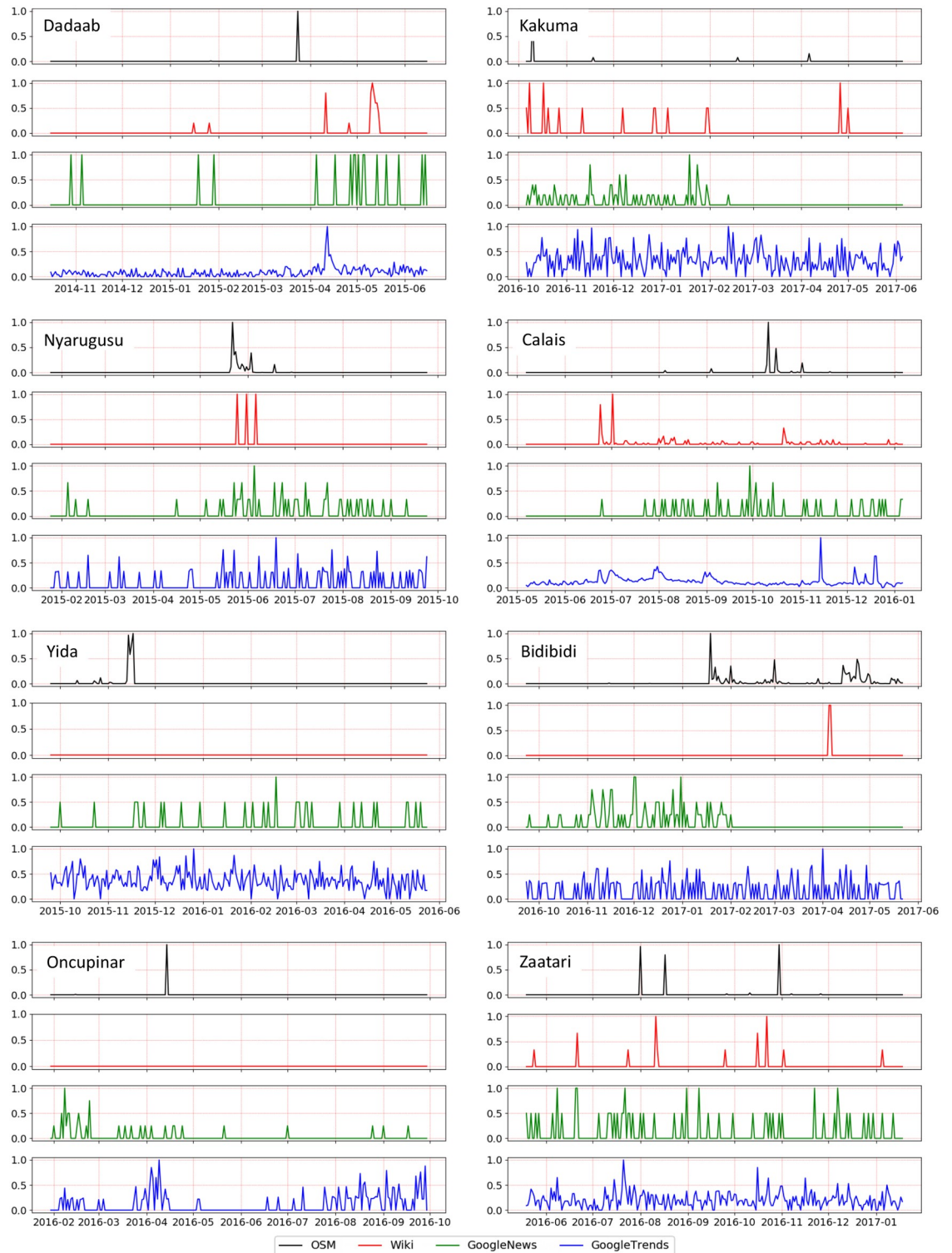
Fig 5. OSM, Wikipedia, Google News, and Google Trends time series during a ±4 months period around the strongest extremum point of each camp. The figures show that whereas OSM and Wikipedia entries tend to come in bursts, Google News and Trends display a more sustained type of activity.

https://doi.org/10.1371/journal.pone.0206825.g005

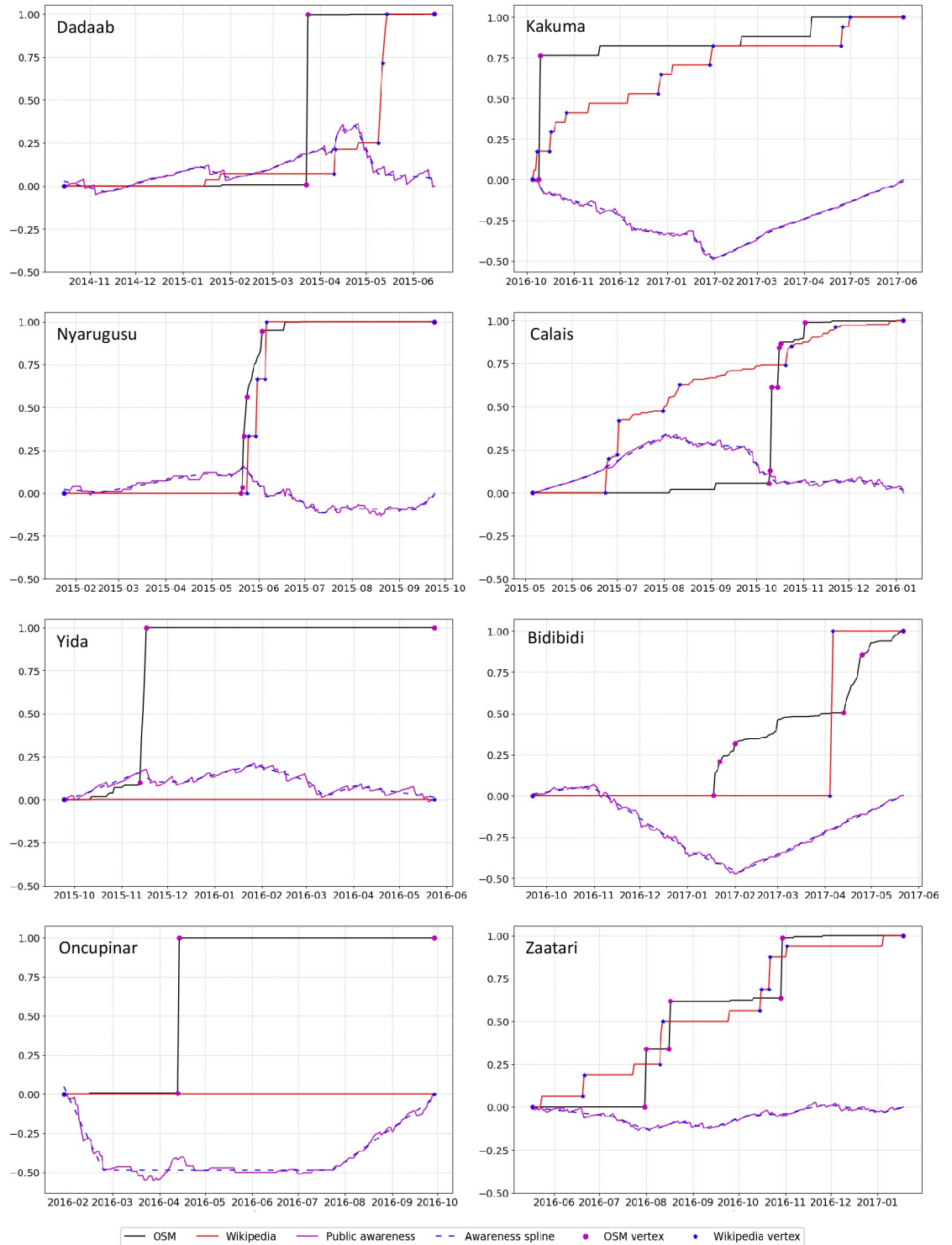**Fig 6. The public awareness curve versus the cumulative OSM and Wikipedia edit activity during a ±4 months period around the strongest extremum point of each camp.** For camps such as Nyarugusu, OSM and Wikipedia bursts overlap with public awareness surplus. In other camps, such as Bidibidi, OSM edit activity bursts coincide with public awareness deficit.

https://doi.org/10.1371/journal.pone.0206825.g006

Table 4. Summary of the time gap values derived at the monthly, weekly, and daily levels for OSM and Wikipedia.

Platform	Time granularity units	Average time gap (and range)
OSM	Monthly	104.6 (0, 334)
	Weekly	43.0 (0, 126)
	Daily	10.8 (1, 32)
Wikipedia	Monthly	108.8 (0, 304)
	Weekly	44.3 (0, 133)
	Daily	12.0 (1, 27)

https://doi.org/10.1371/journal.pone.0206825.t004

campaigns (e.g. [62]), environment (e.g. [111]) and climate change (e.g. [112]) issues. However, little is still known about the impact that news media coverage has on digital activism, especially as it relates to online crowdsourcing platforms such as OSM and Wikipedia.

Our objective in this paper was to advance our understanding of the complex interrelationships that link media coverage and digital activism by focusing in particular on the geographical dimension of news media and the manifestation of digital activism—in particular edit activity—in both geographic and non-geographic crowdsourcing platforms. In order to pursue this goal, we used refugee camps as a test case. As noted earlier, refugee camp sites are particularly suitable for this objective due to their conceptual alignment to the motivational factors that drive volunteerism, and due to their distinct geographical locations, that render their OSM and Wikipedia contributions exclusively related to the camps themselves. Our analysis focused on two interrelated themes, namely the saliency of refugee camps as geographically distinct subjects of digital activism, and the possible co-occurrences between public awareness trends and digital activism. In both cases edit activity of contributors in OSM and Wikipedia was considered to be a manifestation of digital activism.

Considering the issue of saliency, we examined the OSM edit activity in each camp site and a set of three surrounding zones, both in terms of the number of edits and in terms of unique contributors. It was found that, in general, both the total number of edits and the number of edits per km^2 drops substantially around camp sites compared to the edit activity within them. Additionally, a similar decay was found with respect to the number of unique contributors that were engaged in edit activities in camp sites versus the zones surrounding these sites. These results suggest that refugee camp sites tend to serve as geographically salient features that attract purposeful digital activism. Moreover, the decay in the number of unique contributors around the periphery of the camp sites suggest that the camp sites become salient objects of awareness to which OSM contributors pay specific attention. These findings give rise to the idea that the geographic saliency and awareness saliency are interdependent in the context of digital activism.

Focusing on the notion of awareness saliency and digital activism, we then explored the relationship between public awareness and evidence of digital activism related to refugee camps. Using a public awareness measure that was derived from Google News and Google Trends, we compared trend changes in public awareness to patterns of edit activity in the OSM and Wikipedia crowdsourcing platforms. Our findings indicate that in these platforms digital activism bursts tend to take place during periods of build-up of public awareness surplus or deficit, with an average time gap of approximately 11 to 12 days from extremum points in OSM and Wikipedia activity curves to the closest extremum point in the public awareness curve. It is important to note that the average time gap values were consistent across the two crowdsourcing platforms for all tree time granularities that were examined (namely monthly, weekly, and daily). However, our analysis shows that these two platforms do not always share similar activity patterns. Specifically, the results suggest that OSM edit activity within refugee

camps tends to be concentrated in distinct bursts, while Wikipedia edit activity is often characterized by a gradual sustained edit activity effort. This difference suggests that while the user communities in both platforms are potentially exposed to the same public awareness trends, the response of each community may not be the same.

While our analysis did not address directly the issue of motivation, the results of our public awareness analysis highlight the multifaceted nature of motivation in the context of crowdsourcing. Specifically, our results indicated that edit activity bursts can occur during periods of sustained surplus or deficit in public awareness. These seemingly contradicting findings can be explained by two complementary theories in mass communication related to activism, namely agenda setting theory [99] and corrective action theory [113]. A manifestation of the former is the finding that periods of consistent public awareness surplus lead to increased saliency of the corresponding refugee camps as attention artifacts, which in turn lead to edit activity bursts (as is the case with Dadaab, Nyarugusu, and Yida). A manifestation of the latter is the finding that periods of consistent public awareness deficit (as is the case with the rest of the camps), which increases the saliency of the camp sites as attention artifacts due to the perceived lack of coverage of the topic in the news media, and leads to edit activity.

Combined, these results suggest the potential of a novel stimulus-awareness-activism (SA^2) framework in today's participatory digital age. This framework, as presented in Fig 7, is built on three primary constructs: (1) stimulus, (2) awareness, and (3) activism. In this framework, *stimulus* is provided by news media coverage of a specific topic (expressed through Google News metrics). We argue that over time, such coverage leads to *awareness*, whereby the public seeks additional information on the topic (expressed through Google Trends metrics). When awareness grows faster than news coverage a build-up of awareness surplus occurs: a topic resonates with the public, and in a sense goes viral. When awareness growth lags in comparison to news coverage awareness deficits occurs: a topic fails to capture the public's interest and slowly fades away. Build-ups of awareness surplus or deficits lead to activism. Such activism can be manifested either offline (e.g. participating in crowdfunding efforts or volunteering for a non-government organization) or online (e.g. participating in OSM or Wikipedia activities). One could reasonably expect that this process is cyclical in nature, as activism is likely to lead to increased news coverage, providing renewed stimulus and creating a feedback loop in this stimulus-awareness-activism (SA^2) framework. In terms of the individuals involved in such activities, it's important to note that while a large population of individuals may be exposed to the stimulus, only a portion of this population may develop awareness, and an even smaller portion will engage in activism. Additionally, it is important to point out that new individuals may become exposed and develop awareness as the topic gains saliency in the news awareness ecosystem.

In this paper, we studied the framework that links news coverage to awareness and activism by focusing in particular on the question of geographical saliency. Certain news stories tend to have a specific geographic dimension associated with them, and refugee camps are an excellent example of this, both explicitly and implicitly. *Explicitly* because these camps are physical constructs and have a specific location that they occupy (and in some cases are even named after that location), while *implicitly* because they are stops along geographical pathways that take the refugees from an origin location (e.g. the homeland that they had to abandon) to a destination location (their intended final destination). Accordingly, geographical saliency is rather prominent when considering these types of stories. However, geography is also prominent in most other news stories [114], just like for example geographical content is prominent in the vast percentage of Wikipedia entries [115].

As with any study, this study has some noteworthy limitations. Here, we highlight several such limitations that could be further investigated and refined. The first relates to the number

Fig 7. Stimulus-awareness-activism (SA^2) framework.

https://doi.org/10.1371/journal.pone.0206825.g007

of refugee camps studied. In our paper, although we selected refugee camps from around the world, only 8 camps were used. However, a larger cross-section of camps would be useful in exploring the relationship in news media coverage and digital geo-activism in greater breadth. Such research can also benefit from a much longer-term study of these variables, which can further be used to better understand the movement of these variables overtime and their

possible association with other exogenous factors (e.g. crisis events). Second, while our analysis was done in the context of refugee camps, further analysis is required in order to explore whether similar patterns can be observed in other contexts (e.g. natural and manmade disasters, and disease outbreaks). Third, as noted above, the examination of motivational factors is not a straightforward process and requires further investigation (e.g. large-scale surveys of the motivational factors of OSM contributors).

Another, fourth, issue relates to the focus of this study only on the news awareness ecosystem in the English language, primarily due to the overwhelming pervasiveness of English in online content compared to other languages [116]. As the behavior of the news awareness ecosystem in other languages may be different, further analysis of the possible relationship between public awareness and digital activism across different languages is needed. A fifth issue we highlight concerns the use of the volume of Google News items in the public awareness measure used in this work. By considering only the volume and not the content or the impact of each news item our approach takes a simplistic view in which all news items are regarded as equal. However, in practice it is possible that some news items (or news outlets) may become more influential than others, which may result in a different pattern of digital activism. Examining this issue requires a separate line of inquiry that involves the development of appropriate measures for estimating the influence of news items as well as content analysis.

A sixth related issue is that our study only considers extreme bursts of activity in public awareness that tend to trigger digital geo-activism. However, the input variables used in determining such public awareness are expected to be in a perpetual state of fluctuation, with their own unique circadian patterns and influenced by various factors such as seasonality and crisis situations, among others. A more in-depth study analyzing these specific patterns would therefore be of interest. Another related limitation of our study is that for some camps, namely Yida and Oncupinar, there were no Wikipedia edits for these camps for the specific search window used when examined at the daily granular level. In the case of Oncupinar, recent reports have shown that Wikipedia editing access in Turkey was blocked by government authorities in April 2017 [117]. This period, however, is beyond our study period. Nevertheless, it is difficult to assess the impact that such actions may have on inactive periods of edit activity in online platforms such as Wikipedia since for example, technology measures exist that may nullify their effect (e.g. [118–119]). Recent studies, for example, have shown that digital censorship may also have the opposite effect, that is, there is an increase drive to access more information, and thus innovative ways to overcome such restrictions emerge (e.g. [120]).

Further work could also address the degree to which geographical saliency drives digital activism in news stories that relate to geographical areas that are not as monothematic as refugee camps, but rather are often featured in news stories for a wide variety of issues. For example, a megacity may find itself in the news following a major disaster, yet at the same time, it may also be featured for the numerous other activities/issues that are associated with it. Studying such multi-thematic geographical areas will allows us to further refine the SA^2 framework. Such refinements will allow us to devise more effective communication campaigns that will harness the power of the crowd in an organized manner to build responses to societal needs, such as mapping uncharted parts of the Sub-Saharan Africa to better study the birth and spread of exotic diseases at the human-environment interface. Even more importantly, such studies will offer us a better understanding of how our societies function across the cyber-physical news awareness ecosystem that is becoming the prevailing paradigm when interacting with the news.

Supporting information

S1 File. Population references for refugee camps.
(DOCX)

S2 File. OSM edits, Wikipedia edits and Google News articles at the daily level for the period 01/01/2010 to 05/31/2017.
(DOCX)

S3 File. OSM edits within camps and surrounding areas (±4 month period) around the strongest extremum point extracted from the weekly data at the daily granularity.
(DOCX)

Acknowledgments

Author Contributions

Conceptualization: Ron Mahabir, Arie Croitoru, Andrew Crooks, Peggy Agouris, Anthony Stefanidis.

Data curation: Ron Mahabir.

Formal analysis: Ron Mahabir.

Methodology: Ron Mahabir, Arie Croitoru, Andrew Crooks, Peggy Agouris, Anthony Stefanidis.

Writing – original draft: Ron Mahabir.

Writing – review & editing: Ron Mahabir, Arie Croitoru, Andrew Crooks, Peggy Agouris, Anthony Stefanidis.

References

1. Pew. New use across social media platforms 2016. Pew Research Center. 2016. http://www.journalism.org/2016/05/26/news-use-across-social-media-platforms-2016

2. Grazian D. A Digital Revolution? A reassessment of new media and cultural production in the digital age. Annals of the American Academy of Political and Social Science. 2005. 597(1): 209–222. https://doi.org/10.1177/0002716204270286

3. Batsell J. Engaged journalism: Connecting with digitally empowered news audiences. 1st ed. New York: Columbia University Press. 2015

4. Haim M, Graefe A., Brosius H. Burst of the filter bubble. Journal of Digital Journalism. 2017. pp. 1–14. https://doi.org/10.1080/21670811.2017.1338145

5. Redden J, Witschge T. A new news order? Online news content examined. In: Fenton N, editor. New media, old news: Journalism and democracy in the digital age. London: Sage Publications Limited; 2010. pp.171–186.

6. Nielsen RK, Schrøder KC. The relative importance of social media for accessing, finding, and engaging with news: an eight-country cross-media comparison. Digital journalism. 2014. 2(4), pp.472–489. https://doi.org/10.1080/21670811.2013.872420

7. Allcott H, Gentzkow M. Social media and fake news in the 2016 election. Journal of Economic Perspectives. 2017. 31(2): 211–236. https://doi.org/10.1257/jep.31.2.211

8. Yin J, Lampert A, Cameron M, Robinson B, Power R. Using social media to enhance emergency situation awareness. IEEE Intelligent Systems. 2012. 27(6):52–59. https://doi.org/10.1109/MIS.2012.6

9. González-Bailón S, Wang N. Networked discontent: The anatomy of protest campaigns in social media. Social networks. 2016. 44:95–104. https://doi.org/10.1016/j.socnet.2015.07.003

10. McCombs M. Setting the agenda: Mass media and public opinion. 2nd edition. Cambridge: Polity Press. 2014.

11. Xiong F, Liu Y. Opinion formation on social media: an empirical approach. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2014. 24(1):013130. https://doi.org/10.1063/1.4866011 PMID: 24697392

12. Vonbun-Feldbauer R, Matthes J. Do Channels Matter? Investigating media characteristics in the agenda-building process of an election campaign. Journalism Studies. 2017. pp. 1–20.

13. Edwards F, Howard PN, Joyce M. Digital Activism and Non-Violent Conflict. 2013. SSRN. https://ssrn.com/abstract=2595115

14. Khondker HH. Role of the new media in the Arab Spring. Globalizations. 2011. 8(5): 675–679. https://doi.org/10.1080/14747731.2011.621287

15. Howard PN, Parks MR. Social media and political change: Capacity, constraint, and consequence. Journal of communication. 2012. 62(2):359–362. https://doi.org/10.1111/j.1460-2466.2012.01626.x

16. Bruns A, Highfield T, Burgess J. The Arab Spring and social media audiences: English and Arabic Twitter users and their networks. American Behavioral Scientist. 2013. 57(7):871–898.

17. Croitoru A, Wayant N, Crooks A, Radzikowski J, Stefanidis A. Linking cyber and physical spaces through community detection and clustering in social media feeds. Computers, Environment and Urban Systems. 2015. 53: 47–64. https://doi.org/10.1016/j.compenvurbsys.2014.11.002

18. DeLuca KM, Brunner E, Sun Y. Weibo, WeChat, and the transformative events of environmental activism on China's wild public screens. International Journal of Communication. 2016 10:321–339.

19. Qi H, Manrique P, Johnson D, Restrepo E, Johnson NF. Open source data reveals connection between online and on-street protest activity. EPJ Data Science. 2016. 5:18. https://doi.org/10.1140/epjds/s13688-016-0081-5 PMID: 27471660

20. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. Journal of medical Internet research. 2013. 15(4):e85. https://doi.org/10.2196/jmir.1933 PMID: 23615206

21. De Choudhury M, Morris MR, White, RW. Seeking and sharing health information online: Comparing search engines and social media. Proceedings of the 32nd annual ACM conference on Human factors in computing systems. 2014. pp 1365–1376.

22. Xu B, Li D. An empirical study of the motivations for content contribution and community participation in Wikipedia. Information and management. 2015. 52(3):275–286. https://doi.org/10.1016/j.im.2014.12.003

23. Nov O. What motivates Wikipedians? Communications of the ACM, 2007. 50(11):60–64. https://doi.org/10.1145/1297797.1297798

24. Clary EG, Snyder M, Ridge RD, Copeland J, Stukas AA, Haugen J, et al. Understanding and assessing the motivations of volunteers: a functional approach. Journal of Personality and Social Psychology. 1998. 74(6):1516–1530. https://doi.org/10.1037/0022-3514.74.6.1516 PMID: 9654757

25. Budhathoki NR, Haythornthwaite C. Motivation for open collaboration: Crowd and community models and the case of OpenStreetMap. American Behavioral Scientist. 2013. 57(5):548–575. https://doi.org/10.1177/0002764212469364

26. Haklay M, Weber P. 2008. Openstreetmap: User-generated street maps. IEEE Pervasive Computing. 7(4):12–18. https://doi.org/10.1109/MPRV.2008.80

27. Neis P, Zipf A. Analyzing the contributor activity of a volunteered geographic information project—The case of OpenStreetMap. ISPRS International Journal of Geo-Information. 2012. 1(2):146–165. https://doi.org/10.3390/ijgi1020146

28. Rice M, Paez FI, Mulhollen AP, Shore BM, Caldwell DR. Crowdsourced geospatial data. A report on the emerging phenomena of crowdsourced and user-generated geospatial data. 2012. http://www.dtic.mil/docs/citations/ADA576607

29. Coleman DJ, Georgiadou Y, Labonte J. Volunteered Geographic Information: The Nature and Motivation of Producers. International Journal of Spatial Data Infrastructures Research. 2009. 4(1): 332 358. https://doi.org/10.2902/1725-0463.2009.04.art16

30. Napolitano M, Mooney P. 2012. MVP OSM: A tool to identify areas of high quality contributor activity in OpenStreetMap. The Bulletin of the Society of Cartographers. 2012. 45(1):10–18.

31. Zastro M. Crisis mappers turn to citizen scientists. Crowdsourced disaster surveys strive for more reliability in online collaboration. 2014. Nature. 515 (7527).

32. Westrope C, Banick R, Levine M. Groundtruthing OpenStreetMap building damage assessment. Procedia Engineering. 2014 78: 29–39. https://doi.org/10.1016/j.proeng.2014.07.035

33. Dittus M, Quattrone G, Capra L. Mass participation during emergency response: Event-centric crowd-sourcing in humanitarian mapping. Proceedings of the ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017: 1290–1303.

34. Bégin D, Devillers R, Roche S., Contributors' enrollment in collaborative online communities: the case of OpenStreetMap. Geo-spatial Information Science. 2017. 20(3):282–295. https://doi.org/10.1080/10095020.2017.1370177

35. Steiner T, Van Hooland S, Summers E. MJ no more: using concurrent Wikipedia edit spikes with social network plausibility checks for breaking news detection. Proceedings of the 22nd International Conference on World Wide Web. 2013. pp 1–4.

36. Geiß S, Leidecker M, Roessing T. The interplay between media-for-monitoring and media-for-searching: How news media trigger searches and edits in Wikipedia. New Media and Society. 2016. 18 (11):2740–2759. https://doi.org/10.1177/1461444815600281

37. Georgescu M, Kanhabua N, Krause D, Nejdl W, Siersdorfer S. Extracting event-related information from article updates in Wikipedia. European Conference on Information Retrieval. 2013. pp. 254–266.

38. Ciglan M, Nørvåg K. WikiPop: personalized event detection system based on Wikipedia page view statistics. Proceedings of the 19th ACM international conference on Information and knowledge management. 2010. pp. 1931–1932.

39. Google Trends. 2017. https://trends.google.com/

40. Scharkow M, Vogelgesang J. Measuring the public agenda using search engine queries. International Journal of Public Opinion Research. 2011. 23(1):104–113. https://doi.org/10.1093/ijpor/edq048

41. Troumbis AY. Declining Google Trends of public interest in biodiversity: semantics, statistics or traceability of changing priorities? Biodiversity and Conservation. 2017. 26(6):1495–1505. https://doi.org/10.1007/s10531-017-1294-z

42. Preis T, Moat HS, Stanley HE. Quantifying trading behavior in financial markets using Google Trends. Scientific reports. 2013. 3(1684). https://doi.org/10.1038/srep01684 PMID: 23619126

43. Curme C, Preis T, Stanley HE, Moat HS. Quantifying the semantics of search behavior before stock market moves. Proceedings of the National Academy of Sciences. 2014. 111(32):11600–11605. https://doi.org/10.1073/pnas.1324054111 PMID: 25071193

44. Carneiro HA, Mylonakis E. Google trends: a web-based tool for real-time surveillance of disease outbreaks. Clinical infectious diseases. 2009. 49(10):1557–1564. https://doi.org/10.1086/630200 PMID: 19845471

45. Sudhakar NV, Wayda B, Ranasinghe I, Wang S, Dreyer RP, Chen SI, et al. The use of Google Trends in health care research: A systematic review. PLoS ONE. 2014. 9(10): e109583. https://doi.org/10.1371/journal.pone.0109583 PMID: 25337815

46. Cavazos-Rehg PA, Krauss MJ, Spitznagel EL, Lowery A, Grucza RA, Chaloupka FJ, et al. Monitoring of non-cigarette tobacco use using Google Trends. Tobacco control. 2014. 24(3):249–255. https://doi.org/10.1136/tobaccocontrol-2013-051276 PMID: 24500269

47. Ratkiewicz J, Flammini A, Menczer F. Traffic in social media I: paths through information networks. IEEE Second International Conference on Social Computing. 2010. pp. 452–458.

48. Anderson CW. Media ecosystems: Some notes toward a genealogy of the term and an application of it to journalism research. In Unpublished paper presented at ESF Exploratory Workshop on Mapping the digital news ecosystem: Professional journalism, new producers and active audiences in the digital public sphere. 2013. Vol. 8. https://pdfs.semanticscholar.org/a8e4/edef4a7d1c65ae15254deb0756e402fc4914.pdf

49. Picard RG. Twilight or new dawn of journalism? Evidence from the changing news ecosystem. Digital Journalism. 2014. 2(3): 273–283. https://doi.org/10.1080/1461670X.2014.895530

50. Nghiem LT, Papworth SK, Lim FK, Carrasco LR. Analysis of the capacity of Google Trends to measure interest in conservation topics and the role of online news. PLoS ONE. 2016. 11(3): e0152802. https://doi.org/10.1371/journal.pone.0152802 PMID: 27028399

51. Crooks A, Croitoru A, Stefanidis A, Radzikowski J. # Earthquake: Twitter as a distributed sensor system. Transactions in GIS. 2013. 17(1):124–147. https://doi.org/10.1111/j.1467-9671.2012.01359.x

52. Tandoc EC Jr, Vos TP. The journalist is marketing the news: Social media in the gatekeeping process. Journalism Practice. 2016. 10(8): 950–966. https://doi.org/10.1080/17512786.2015.1087811

53. Althoff T, Borth D, Hees J, Dengel A. Analysis and forecasting of trending topics in online media streams. Proceedings of the 21st ACM international conference on Multimedia. 2013. pp 907–916.

54. Al Emadi N, Leson H, Lucas J, Meier P, Borge-Holthoefer J. Geography of Solidarity: Spatial and Temporal Patterns. Procedia Engineering. 2016. 159:27–34. https://doi.org/10.1016/j.proeng.2016.08.060

55. MicroMappers. 2017. http://clickers.micromappers.org

56. Heinzelman J, Waters C. *Crowdsourcing crisis information in disaster-affected Haiti.* Washington DC: Institute of Peace. 2010.

57. Zook M, Graham M, Shelton T, Gorman S. Volunteered geographic information and crowdsourcing disaster relief: a case study of the Haitian earthquake. World Medical & Health Policy. 2010. 2(2):7–33. https://doi.org/10.2202/1948-4682.1069

58. See L, Comber A, Salk C, Fritz S, van der Velde M, Perger C, et al. Comparing the quality of crowd-sourced data contributed by experts and non-experts. PLoS ONE. 2013. 8(7): e69958. https://doi.org/10.1371/journal.pone.0069958 PMID: 23936126

59. Mullen WF, Jackson SP, Croitoru A, Crooks A, Stefanidis A, Agouris P. Assessing the impact of demographic characteristics on spatial error in volunteered geographic information features. 2017. GeoJournal. 80(4), pp.587–605. https://doi.org/10.1007/s10708-014-9564-8

60. Crooks A, Croitoru A, Jenkins A, Mahabir R, Agouris P, Stefanidis A. User-Generated Big Data and Urban Morphology. Built Environment. 2016. 42, pp.396–414. https://doi.org/10.2148/benv.42.3.396

61. Ku G, Kaid LL, Pfau M. The impact of web site campaigning on traditional news media and public information processing. Journalism & Mass Communication Quarterly. 2003. 80(3):528–547.

62. Lee B, Lancendorfer KM, Lee KJ. 2005. Agenda-setting and the Internet: The intermedia influence of Internet bulletin boards on newspaper coverage of the 2000 general election in South Korea. Asian Journal of Communication. 15(1), pp.57–71. https://doi.org/10.1080/0129298042000329793

63. Messner M, South J. Legitimizing Wikipedia: How US national newspapers frame and use the online encyclopedia in their coverage. Journalism Practice. 2011 5(2):145–160. https://doi.org/10.1080/17512786.2010.506060

64. Crooks A, Pfoser D, Jenkins A, Croitoru A, Stefanidis A, Smith D, et al. Crowdsourcing urban form and function. International Journal of Geographical Information Science. 2015. 29(5), pp.720–741. https://doi.org/10.1080/13658816.2014.977905

65. Mahabir R, Stefanidis A, Croitoru A, Crooks AT, Agouris P. Authoritative and volunteered geographical information in a developing country: A comparative case study of road datasets in Nairobi, Kenya. ISPRS International Journal of Geo-Information. 2017. 6(1):24. https://doi.org/10.3390/ijgi6010024

66. Berry M, Garcia-Blanco I, Moore K. Press coverage of the refugee and migrant crisis in the EU: a content analysis of five European countries. 2015. United Nations High Commission for Refugees (UNHCR). http://www.unhcr.org/56bb369c9.html

67. World Bank. Forcibly displaced: Towards a development approach supporting refugees, the internally displaced, and their hosts. 2017. https://openknowledge.worldbank.org/handle/10986/25016

68. Koren M. Where are the 50 most populous refugee camps? Smithsonian Magazine. 2013. http://www.smithsonianmag.com/innovation/where-are-50-most-populous-refugee-camps-180947916/

69. OpenStreetMap. Stats. 2018. http://wiki.openstreetmap.org/wiki/Stats

70. Doctors without Borders 2015. GIS Support for the MSF Ebola Response in Guinea in 2014. http://www.cartong.org/sites/cartong/files/GIS%20Support%20Ebola%202015_EN.pdf

71. Rehrl K, Gröchenig S. A framework for data-centric analysis of mapping activity in the context of volunteered geographic information. International Journal of Geo-Information. 2016. 5(3): 37. https://doi.org/10.3390/ijgi5030037

72. Hristova D, Mashhadi A, Quattrone G, Capra L. Mapping community engagement with urban crowdsourcing. AAAI Technical Report WS-12-04. 2012. https://pdfs.semanticscholar.org/94f2/d7233ed7ceaeb396a4871431dd66de9f4184.pdf

73. Ma D, Sandberg M, Jiang B. Characterizing the heterogeneity of the OpenStreetMap data and community. International Journal of Geo-Information. 2015. 4(2): 535–550. https://doi.org/10.3390/ijgi4020535

74. Anthony D, Smith SW, Williamson T. The quality of open source production: Zealots and good samaritans in the case of Wikipedia. Rationality and Society. 2007. pp 32

75. Buriol L, Castillo C, Donato D, Leonardi S, Millozzi S. Temporal evolution of the wikigraph. Proceedings of the International Conference on Web Intelligence. 2006. pp 45–51.

76. Java A, Song X, Finin T, Tseng B. Why we twitter: understanding microblogging usage and communities. Proceedings of the 9th WebKDD and 1st SNA-KDD. 2007 workshop on Web mining and social network analysis. pp. 56–65.

77. Zhang Y, Sun A, Datta A, Chang K, Lim EP. Do Wikipedians follow domain experts?: a domain-specific study on Wikipedia knowledge building. Proceedings of the 10th annual joint conference on Digital libraries. 2010. pp. 119–128.

78. Wilkinson DM, Huberman BA. Assessing the value of cooperation in Wikipedia. First Monday. 2007. 12(4). Available from: http://firstmonday.org/article/view/1763/1643

79. Wattenberg M, Viégas FB, Hollenbach K. Visualizing activity on wikipedia with chromograms. IFIP Conference on Human-Computer Interaction. 2007. pp 272–287.

80. Vandecasteele A, Devillers R. Improving volunteered geographic information quality using a tag recommender system: the case of OpenStreetMap. In: OpenStreetMap in GIScience. Arsanjani JJ, Zipf A, Mooney P, Helbich M. (editors). Springer International Publishing. 2015. pp 59–80.

81. Haklay M. 2010. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and planning B: Planning and design, 37 (4):682–703. https://doi.org/10.1068/b35097

82. Sui D, Goodchild M, Elwood S. Volunteered geographic information, the exaflood, and the growing digital divide. In: Crowdsourcing geographic knowledge. Sui D, Goodchild M, Elwood S. (editors). 2013. Springer: Netherlands. pp. 1–12.

83. Barrington-Leigh C, Millard-Ball A. The world's user-generated road map is more than 80% complete. 2017, PLoS ONE. 12(8): e0180698. https://doi.org/10.1371/journal.pone.0180698 PMID: 28797037

84. Mooney P, Corcoran P. April. How social is OpenStreetMap. Proceedings of the 15th association of geographic information laboratories for Europe international conference on geographic information science. 2012. pp 24–27

85. Mooney P, Corcoran P. Analysis of Interaction and o-editing patterns amongst OpenStreetMap Contributors. Transactions in GIS. 2014. 18(5):633–659. https://doi.org/10.1111/tgis.12051

86. Hristova D, Quattrone G, Mashhadi AJ, Capra L. The Life of the Party: Impact of Social Mapping in OpenStreetMap. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media. 2013. pp 234–243.

87. Mashhadi A, Quattrone G, Capra L. The impact of society on volunteered geographic information: The case of OpenStreetMap. In: OpenStreetMap in GIScience. Arsanjani JJ, Zipf A, Mooney P, Helbich M. (editors). Springer International Publishing. 2015. pp. 125–141.

88. Yang A, Fan H, Jing N, Sun Y, Zipf A. Temporal analysis on contribution inequality in OpenStreetMap: A comparative study for four countries. International Journal of Geo-Information. 2016. 5(1). pp 5. https://doi.org/10.3390/ijgi5010005

89. Wikipedia. Wikipedia: Statistics. 2018. https://en.wikipedia.org/wiki/Wikipedia:Statistics

90. Google News. 2017. https://news.google.com/

91. Bharat. Google turns 10. 2012 September. [cited 15 September 2017] In: Google Journalism and News Blogs. https://www.blog.google/topics/journalism-news/google-news-turns-10/

92. Weaver DA, Bimber B. Finding news stories: a comparison of searches using LexisNexis and Google News. Journalism and Mass Communication Quarterly. 2008. 85(3):515–530. https://doi.org/10.1177/107769900808500303

93. Granka L. Measuring agenda setting with online search traffic: influences of online and traditional media. Annual Meeting of the American Political Science Association. 2010. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1658172

94. Ripberger JT. Capturing curiosity: Using Internet search trends to measure public attentiveness. Policy Studies Journal. 2011. 39(2), 239–259. https://doi.org/10.1111/j.1541-0072.2011.00406.x

95. The Cambridge Dictionary. 2018. https://dictionary.cambridge.org/us/dictionary/english/awareness

96. Zhang W, Goh ATC. Multivariate adaptive regression splines and neural network models for prediction of pile drivability. Geoscience Frontiers. 2016. 7(1) 45:52. https://doi.org/10.1016/j.gsf.2014.10.003

97. Friedman JH. Multivariate Adaptive Regression Splines. The Annals of Statistics. 1991. 19:1.

98. Stone GC, McCombs ME. Tracing the time lag in agenda-setting. Journalism Quarterly. 1981. 58(1):51–55.

99. McCombs M, Shaw D. The agenda-setting function of mass media. Public Opinion Quarterly. 1972. 36 (2): 176. https://doi.org/10.1086/267990

100. Roberts M, Wanta W, Dzwo TH. Agenda setting and issue salience online. Communication Research. 2002. 29(4):452–465. https://doi.org/10.1177/0093650202029004004

101. Shoemaker PJ, Wanta W, Leggett D. Drug coverage and public opinion. In: Shoemaker PJ. (ed.), Communication campaigns about drugs: Government, media, and the public Hillsdale, NJ: Lawrence Erlbaum. 1989. pp 67–80.

102. Douglas D, Peucker T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. The Canadian Cartographer. 1973. 10(2):112–122. https://doi.org/10.3138/FM57-6770-U75U-7727

103. Barabasi A. Bursts: The hidden patterns behind everything we do, from your email to bloody crusades. Reprint edition. New York: Penguin Group. 2011.

104. Purcell K, Rainie L, Mitchell A, Rosenstiel T. and Olmstead K. Understanding the participatory news consumer. Pew Internet and American Life Project, 1, 2010. pp 19–21. Available from: http://www.journalism.org/files/legacy/Participatory_News_Consumer.pdf

105. Deuze M. Participation, remediation, bricolage: Considering principal components of a digital culture. The information society. 2006. 22(2):63–75. https://doi.org/10.1080/01972240600567170

106. Mitchelstein E, Boczkowski PJ. Between tradition and change: A review of recent research on online news production. Journalism. 2009. 10(5):562–586. https://doi.org/10.1177/1464884909106533

107. Joyce MC. Digital activism decoded: The new mechanics of change. IDEA. 2010. Brussels, Belgium: International Debate Education Association.

108. Earl J, Kimport K. Digitally enabled social change: Activism in the internet age. 2011. Boston, Massachusetts: MIT Press.

109. Sayre B, Bode L, Shah D, Wilcox D, Shah C. Agenda setting in a digital age: Tracking attention to California Proposition 8 in social media, online news and conventional news. Policy and Internet. 2010. 2(2):7–32. https://doi.org/10.2202/1944-2866.1040

110. Benkler Y, Roberts H, Faris R, Solow-Niederman A, Etling B. Social mobilization and the networked public sphere: Mapping the SOPA-PIPA debate. Political Communication. 2015. 32(4): 594–624. https://doi.org/10.1080/10584609.2014.986349

111. Lester L, Hutchins B. Power games: environmental protest, news media and the internet. Media, Culture and Society. 2009. 31(4): 579–595. https://doi.org/10.1177/0163443709335201

112. O'Neil S, Boykoff Maxwell. The role of new media in engaging the public with climate change. Chapter 13. In Engaging the Public with Climate Change: Behaviour Change and Communication. Whitmarsh L, O'Neill S, Lorenzoni. (eds). 2011. Washington, DC: Earthscan.

113. Rojas H, Puig-i-Abril E. Mobilizers mobilized: Information, expression, mobilization and participation in the digital age. Journal of Computer-Mediated Communication. 2009. 14(4):902–927.

114. Leetaru K, Schrodt PA. GDELT: Global data on events, location, and tone. In ISA Annual Convention. 2013.

115. Lehmann J, Isele R, Jakob M, Jentzsch A, Kontokostas D, Mendes PN, et al. DBpedia—a large-scale, multilingual knowledge base extracted from Wikipedia. Semantic Web. 2015 6(2), pp.167–195.

116. Wikipedia (2018). Languages used on the internet. https://en.wikipedia.org/wiki/Languages_used_on_the_Internet#cite_note-9

117. Wikipedia (2018). Blocking of Wikipedia in Turkey. https://en.wikipedia.org/wiki/Block_of_Wikipedia_in_Turkey

118. Nobori D, Shinjo Y. VPN Gate: A volunteer-organized public VPN relay system with blocking resistance for bypassing government censorship firewalls. In USENIX Symposium on Networked Systems Design and Implementation. 2014. pp. 229–241.

119. Di Florio A, Verde NV, Villani A, Vitali D, Mancini LV. Bypassing censorship: A proven tool against the recent internet censorship in Turkey. In Proceedings of IEEE International Symposium on Software Reliability Engineering Workshops. 2014. pp. 389–394.

120. Hobbs WR, Roberts ME. How sudden censorship can increase access to information. American Political Science Review. 2018. 112(3):621–636. https://doi.org/10.1017/S0003055418000084