Social Multimedia, Diversity, and Global South Cities: A Double Blind Side

Daniel Gatica-Perez
Idiap Research Institute and EPFL, Switzerland
gatica@idiap.ch

Darshan Santani
Pepper Cloud, Singapore
dsantani@peppercloud.com

Joan Isaac-Biel
Amaris, Switzerland
joanisaac.biel@gmail.com

Thanh-Trung Phan
Idiap Research Institute and EPFL, Switzerland
tphan@idiap.ch

ABSTRACT

Social media provides opportunities to examine urban phenomena at scale, and we believe that studying cities in the Global South through citizen-contributed data and AI-driven analytics should be a priority of multimedia research. However, little work has been done in our community, and we argue that this contributes to a double blind side problem. We exemplify this situation by studying Ma3Route, a mobile social media channel to crowdsource and broadcast transit reports in Nairobi, Kenya. Using multimedia data from its Twitter stream, we first conduct a descriptive analysis that shows an active community generating rich traffic-related reports, and then discover latent topics that identify both regular and ephemeral thematic clusters of reports involving accidents, traffic conditions, and attitudes of citizens towards authorities. In the second place, we conduct a deep learning-based analysis of Ma3Route images to understand the kind of visual content shared in the platform, and that shows limitations of using deep neural network models trained with data largely coming from the US and Europe, which do not fully match the reality and diversity of other world regions. We conclude by presenting a multidisciplinary research agenda for future work in this domain.

CCS CONCEPTS

• **Human-centered computing → Ubiquitous and mobile computing**; *Ubiquitous and mobile computing design and evaluation methods.*

KEYWORDS

Social Media; Mobile Crowdsourcing; Cities; Urban computing

ACM Reference Format:
Daniel Gatica-Perez, Darshan Santani, Joan Isaac-Biel, and Thanh-Trung Phan. 2019. Social Multimedia, Diversity, and Global South Cities: A Double Blind Side. In *1st International Workshop on Fairness, Accountability, and Transparency in MultiMedia (FAT/MM '19), October 25, 2019, Nice, France.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3347447.3356749

FAT/MM '19, October 25, 2019, Nice, France

ACM ISBN 978-1-4503-6915-2/19/10…$15.00
https://doi.org/10.1145/3347447.3356749

1 INTRODUCTION

Social media channels are used by citizens across the globe to voice their opinions on social and economical issues faced by the cities they inhabit. According to official statistics, 25 of the 30 largest urban areas in the world are located in countries in the so-called Global South [49]. This would suggest that studying phenomena in large cities through multimedia analytics should include the study of cities in Africa, Asia, and Latin America.

Surprisingly, however, this is not the case for urban multimedia if one examines the body of research in the past years (with the notable exception of China). A review of ACM MM long papers published in the last years shows little or no work on multimedia analytics involving cities in these world regions. In contrast, analytics of multimedia data coming from economically developed cities like New York, London, San Francisco, and others have received substantial attention in the literature [23, 30, 34, 35, 37]. Recent papers specifically published in ACM MM connecting multimedia analysis and such cities include [36] (New York and Boston) and [22] (Rome and Milan).

Our position is that studying cities in the Global South through citizen-contributed data and AI-driven multimedia analytics should be a priority of multimedia research, and that not doing it contributes (by omission) to produce a *double blind side*: the first one with respect to our understanding, as a research community, of phenomena in Global South cities that have different dynamics and issues to solve than wealthy countries; the second one with respect to what machines can learn (or not) about these cities, when they are represented by data that does not reflect the world's diversity.

Cities in the Global South have been studied outside the multimedia literature, often in the realm of development studies [52] and social sciences [28], and frequently focused on the study of crises [43] and violence [32]. Needless to say, most Global South cities are not in a permanent state of crisis [47]. We argue that the study of everyday urban life phenomena in African, Asian, and Latin American cities with multimedia analytics is important because the fact that most of the urban world is in the Global South opens a wealth of opportunities (as many of these cities produce both social media and other citizen-contributed data) to produce valuable findings that could support communities, organizations, and government to address some of the issues citizens face. Along with it, however, come a number of challenges, related to the need to understand the social practices and cultural contexts in which citizen-contributed multimedia is created, shared, and appropriated.

We exemplify this situation with Ma3Route, a platform that uses mobile and social media to crowdsource transit reports in Nairobi, Kenya. It has been estimated that this city loses each year an equivalent of 200 million USD in lost productivity as a result of traffic [8]. Ma3Route provides a community-driven channel that asks citizens to share traffic reports, and then broadcasts these reports using Twitter to inform commuters in real-time. Users also use the platform to voice their opinions on transportation and safety policies, as well as mistrust and frustration towards the city administration and the police. As of July 2019, the platform has 1.14M followers on Twitter.

This case study is relevant for several reasons. First, by investigating the Ma3Route Twitter channel, we contribute to the examination of the urban dynamics and mobility in Global South cities as expressed in social media, which are often different in characteristics when compared to cities in the US and Western Europe. Second, by studying Nairobi, we analyze a number of common trends that apply to many large urban areas in the world. Nairobi is one of the top 100 populated urban areas worldwide, 80% of which are in the Global South [49]. Finally, the data reveals current limitations of deep learning systems with respect to the diversity of urban scenes represented – an issue that the fairness, accountability and transparency emerging literature has demonstrated for individual citizens in visual tasks like face recognition [20], but that is also applicable to urban areas as we show here,.

The contributions in this paper are three-fold. First, we collected tweets from Ma3Route's Twitter multimedia stream over a period of three months; conduct a descriptive data analysis enriched by qualitative observations made on both tweets and users, which reveals the main characteristics of the platform and the kind of contributed content; and apply topic modeling to discover the main themes of conversation, identifying both regular and ephemeral thematic clusters of citizen contributions, which go beyond accident and traffic reports and reflect the concerns of citizens towards road conditions, transportation, and city authorities. Second, we present a deep-learning based analysis of Ma3Route images to understand the visual content shared in the platform, revealing the diversity limitations of using pretrained deep networks trained with data collected in the US and Western Europe. Third, based on our findings, we propose a agenda for future work to address the double blind side situation.

2 RELATED WORK

2.1 Social Multimedia in the Global South

Ethnographic work examined Facebook usage of young adults living in an informal settlement in Nairobi [51], finding that it is actively used as a medium to search for employment, along with its prototypical use of keeping in touch. A similar Facebook study was conducted among urban youth from disadvantaged communities in urban India [31]. Twitter data was used to analyze national elections of three African nations: Kenya, Ghana and Nigeria [18], finding that in Nigeria, discussions of tribe identity and religion dominate the discourse, while in Ghana citizens were more likely to engage in policy discussions including education. A similar study analyzed the use of Twitter by political parties in the Indian elections [29]. Twitter has been studied to gather eyewitness reports of drug-related violence in Mexico [32], to examine the 2013 Westgate mall terrorist attacks in Nairobi [43], to mobilize people during the Arab Spring [28], and to stimulate debate on sexual violence in Nigeria [17]. Social media for crises is a topic in itself [24]. In contrast to emergency events, we are interested in understanding everyday life phenomena, which have been less studied.

2.2 Mobile Citizen Reporting and Traffic

Crowdsourced reporting of civic issues that follows a "by the people, for the people" philosophy has gained popularity in the recent past. There are many systems that allow the use of mobile and web applications to enable citizens to report local, non-emergency civic issues like reporting potholes, broken street lights, unattended litter, etc. FixMyStreet [9] in the UK and SeeClickFix [12] in the US are two such platforms. FixMyStreet has received hundreds of thousands of reports [44]. Moreover, the local administration in various US cities have released mobile applications to report such issues e.g., Boston has released a mobile application called Bos:311 [3] to help citizens report issues related to city services. In contrast, research in African cities has investigated the use of mobile crowdsourcing to gather and document road quality information in Nairobi [40] or to map the informal bus transportation network [50]. Other campaigns to report urban concerns in Latin America include e.g. [39].

The idea of crowdsourcing traffic updates and broadcasting them to commuters is not new. Waze, which was acquired by Google in 2013, is a popular mobile application with millions of users worldwide [15]. Waze works by collecting GPS sensor data from its users in order to provide real-time traffic updates. As Ma3Route, other similar systems have been created in other countries. In India, Trafline [14] is a mobile and web application, that provides traffic information across several major metropolitan areas. TrafficDito in Manila (Philippines) [13], and ma2too3a in Lebanon [10] were created to provide similar services to their residents. Unfortunately, all these channels, which serve large cities in the Global South, have remained invisible to multimedia research until now.

3 BLIND SIDE #1: MA3ROUTE AS CASE STUDY

We first introduce Ma3Route as a sociotechnical system that illustrates the diversity of issues of Global South cities. We then describe the collected dataset, followed by a data-driven descriptive analysis, and a more in-depth analysis based on topic models. This section reveals that cities in the Global South can indeed be analyzed using citizen-contributed data, but that this promising domain has been largely unaddressed by multimedia research.

3.1 Platform Overview

Ma3Route crowdsources transit reports in Nairobi to provide users with information about traffic, matatu (informal buses in Kenya) directions, and driving reports in almost real-time. The platform aggregates, curates, and broadcasts citizen reports using Twitter, and its own web and mobile applications. Reports consist mainly of text and images, and do not rely on GPS or maps to geolocalize information. Instead, users use road names and other well known places and landmarks as text to locate their reports. This allows

Type of report: Categories
Traffic reports: accident, bad terrain, bumper to bumper, clear, country wide news, general info, impassable, moderate, poor visibility, stationary object, under construction
Driving reports: dangerous, general info, in an accident, involved in crime, well driven

Table 1: Ma3route categories of traffic and driving reports.

Tweet Entity	% Tweets	Total Freq	Num Unique
User mentions	14%	10225	2072
Hashtags	14%	10299	4141
Photos	18%	14138	14023

Table 2: Multimedia entities contained in the dataset.

to overcome technical limitations of mobile devices, poor internet connectivity, or lack of familiarity with western-style maps [7].

Users can create and navigate reports using Ma3Route's mobile or web apps [11]. In both cases, users can create traffic or driving reports by providing a text description and a report category label (Table 1). A second option is to use Twitter, creating reports by simply including the *@Ma3Route* handle in their tweets [2]. The platform aggregates and broadcasts them to their mobile app, web app, and Twitter. Ma3Route uses a convention to acknowledge the author of every tweet by terminating the curated tweets with *"via @username"*. For example, suppose a user writes the following tweet: *"@Ma3Route road traffic so slow"*. Ma3Route's official channel will be automatically updated with the following status: *"road traffic so slow via @username"*. All tweets are categorized as they appear on the apps.

3.2 Ma3Route Dataset

We used the public Twitter REST API to collect tweets posted on the Ma3Route Twitter channel for a period of 3 months (Jun-Sep 2015). This resulted in a collection of 60,086 tweets. A tweet contains the text of the tweet, timestamp, user details, hashtags and media items. In addition, the Ma3Route website shows tweets from its Twitter channel plus associated report categories [11]. The produced Twitter data is relatively small in terms of volume compared to those of large US or European cities. However, we argue that missing the importance of studying urban processes of certain regions because of the relatively modest data volumes they produce is actually part of the blind side problem we discuss here.

3.3 Descriptive Analysis

3.3.1 User contributions. We used a standard regular expression mechanism to process the *"via @username"* part of the tweet to recover the user IDs who authored the original tweet. In the dataset, 97% of tweets were generated by 15,534 unique Twitter users other than Ma3Route itself. The rest of the posts were authored by Ma3Route using their admin handle, or seem to be created by community managers that directly respond or retweet to content using standard Twitter mechanisms, i.e. without the use of the *"via @username"* convention. The number of contributions per user follows a tailed distribution (max=660; median=3.8; mean=1) during the period of our analysis. We inspected the profiles of the top twenty contributors and found local public organizations promoting initiatives on road safety, in addition to Ma3Route itself. The rest of the top twenty contributors are local Nairobians who contributed between 120 and 550 tweets each.

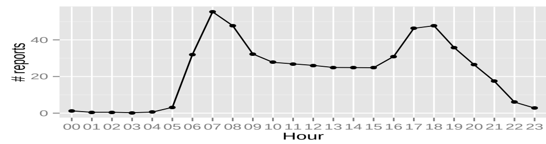

Figure 1: Hourly volume of reports in an average day, aggregated across full dataset, normalized by the number of days.

3.3.2 Multimedia content. We quantified the amount of multimedia content available in terms of user mentions, hashtags, and photos. Table 2 shows the frequencies of these entities in terms of tweets, raw frequency, and unique values.

User mentions: References to users are used to direct content or simply mention users. Twitter notifies users when someone mentions them. Mentions reflect an explicit interest from users towards others, thus looking at the top mentioned users is informative of the content. First, we processed the metadata to avoid counting mentions generated using Ma3Route's *"@via username"* convention, discarding 86% of the tweets and leaving a total of 10,225 user mentions. We inspected the profiles of the top twenty mentioned users and found that most of them were public figures (e.g., Nairobi's Major, Governor Office, etc.) and public institutions (e.g., Kenya's National Transit and Safety Authority, Kenya Police, etc.) responsible for traffic control and security, road infrastructure, and public transport. This shows how the Ma3Route community actively points to the main actors of the city administration.

Hashtags: Hashtags reveal an explicit interest of users to link to specific conversations. We inspected the top 20 hashtags and found that half of them are road names such as *#waiyakiway* or *#mombasaroad*. However, we did not find any evidence of users using road names in hashtags as a general convention to create traffic reports around specific landmarks. We also found other hashtags that refer to road safety campaigns such as *#zushaleo* or *#trafficwatch* or to major national events such as the US president Obama visit to Nairobi in July 2015 (*#obamareturns*, *#kiderograss*), or a matatu incident reported in the media (*#arrestrongaimatatus*).

Photos: Photos complement text reports, and are valuable to document road hazards [40], as testimonial data to investigate accidents, to support road safety campaigns, and to validate the authenticity or reliability of text reports. As shown in Table 2, 18% of tweets contained photos. This is a considerable proportion of photos compared to systems like FixMyStreet that have an estimate of 11% of posts including photos [44].

3.3.3 Reports over time. The time of day of a tweet provides evidence of the situational context that motivates people to report. In Figure 1, we plot the daily evolution of traffic reports as a result of aggregating tweets by time of the day (normalized by the total number of days). The trend peaks in the morning around 7AM and in the evening around 5PM, which corresponds to peak commuting hours in Nairobi. At peak hours, Ma3Route receives between 1–2 reports reports per minute, whereas between these two peaks, the number of reports decreases by half.

Topic	Most Relevant Words
T1	sacco, matatus, driver, embassava, passeng, wrong, bus, shame, reckless, loud, suspend
T2	police, fire, bribe, offic, traffic, cop, fuel, light, car, arrest, collect, motorist, law, corrupt
T3	obama, road, close, mombasa, day, friday, nairobi, time, grass, uhuru, visit, kidero
T4	accid, involv, car, lorri, caus, hit, road, truck, bypass, polic, dead, injury, scene, bus
T5	drive, speed, safeti, cross, pedestrian, #zushaleo, safe, road, limit, life, #trafficwatch, drink
T6	road, langata, jam, mombasa, traffic, ngong, town, happen, bypass, jogoo, msa, park
T7	road, traffic, flow, move, jogoo, slow, citi, inbound, stadium, tow, mbagathi, langata, mombasa
T8	thika, road, muthaiga, move, lane, jam, servic, outer, juja, traffic, kiambu, superhighway, slow
T9	uhuru, highway, avenu, hail, univers, traffic, kenyatta, roundabout, selassi, nyayo, cbd, stuck
T10	waiyaki , westland, kangemi, jame, gichuru, traffic, limuru, kabet, abc, museum, westi, chiromo

Table 3: List of ten discovered topics along with the selection of the most relevant words per topic. Topics are ordered based on their probabilities $P(z)$.

3.3.4 Report categories. We computed a breakdown of the categories assigned by the platform to reports in Table 1. These categories provide information about the reasons that drive the creation of reports. We found that 60% of reports are categorized as *General Info*, while the rest are *Traffic* (*either Bumper to Bumper, Moderate, or Impassable*, 25%), *Accident* (6%), or *Clear* (5.3%). Users tend to report more while stuck in traffic than when the road is clear. This is not surprising, as people might have more time to tweet in dense traffic jams. Given the large proportion of reports labeled as *General Info*, which reduces the information value of this category, we were motivated to use topic models to analyze the content in the reports in more detail, which we present in the next section.

3.4 Topic Analysis

3.4.1 Applying Latent Dirichlet Allocation. Before topic modeling, we filtered and preprocessed the data. We first selected tweets in English. English represents the language of choice amongst Ma3Route users, with 88% of the tweets in English according to Twitter automatic language detection. We then performed standard text preprocessing which included lowercasing, URL, stopword and punctuation removal, and stemming. We also removed user mentions, but kept hashtags as they might potentially contain meaningful keywords for content analysis. We also removed words that appeared in less than 10 tweets. After processing, the dataset consists of 52,626 tweets with a vocabulary size of 3,422 words.

We then used Latent Dirichlet Allocation (LDA) [19]. To train LDA, we treated each tweet as a single document. Previous research has shown that despite the short length of tweets, LDA can obtain meaningful topics from Twitter data [38, 53]. We used a Gibbs sampling implementation [27] requiring two parameters to control the Dirchlet priors on the distribution of topics (α), and the distribution of terms per topic (β). It is common practice to use $\alpha = 50/T$ and $\delta = 0.1$ [26]. We found it appropriate to use a smaller value of alpha α to increase the likelihood of assigning one single topic to each document. Finally, based on a empirical procedure, we found $k = 10$ topics to be suitable for our corpus.

3.4.2 Topic Inspection. In Table 3, we list the ten discovered topics along with their most relevant words. The relevance of the top words is computed using a measure that combines the probability of terms per topic β with the probability of each term in the full corpus [42]. The discovered topics expand our understanding of how citizens use the platform. Topics 6 to 10 are traffic-related topics, while Topics 1 to 5 highlight citizens' concerns towards local issues related to transportation and safety policies. Below, we describe each topic in more depth. Following guidelines for social media research [46], we paraphrase (rather than quote) some of the top tweets for each topic to reflect the opinions of users.

Topic 1 – Matatus ($P(z) = 0.12$). Matatus are privately owned minibuses used for inter and intra-city commute in Nairobi. Due to the lack of public transportation, they provide the necessary transit means to millions of Nairobians every day. Matatus drivers are unfortunately famous for reckless driving, breaking traffic laws, cutting off other vehicles, using sidewalks to bypass traffic jams, and over-speeding [1, 5]. This perception of matatu drivers is reflected in Topic 1 from the top relevant keywords comprising this topic: *matatus*, *drivers*, *passengers*, *suspend*, *loud*, *reckless*, *wrong*, and *shame*. One example complaint is that certain crossing points are rather dangerous, and that pedestrian crossing signals are needed.

Note that both *saccos* and *embassava* are cooperatives of matatu industry. Matatu drivers' income and job security depends on the daily passenger load [25, 33]. In addition, matatus are known for being the target of violence against drivers, bystanders, and commuters. This appears in some of the top tweets, which tell users to beware of specific routes where teams of pickpockets operate.

Topic 2 – Police ($P(z) = 0.12$). This topic reflects the attitude and mistrust of citizens towards police authorities. The relevant words in this topic are *police*, *speed*, *traffic*, *arrest* and *bribe*. In the top tweets, users argue that the city's video camera network might be used by the police for bribing purposes, or complain that the NTSA (the National Transport and Safety Authority, a government body tasked to manage road transport) does not do a proper job.

Topic 3 – Obama ($P(z) = 0.11$). This topic corresponds to the state visit of the U.S. President to Kenya in July 2015. Road closures and disruptions by the city administration along the five major road segments [6] prompted citizens to voice their discomfort to their daily commute and to share traffic updates on road closures as a result of the visit. This is captured in the top terms – *obama*, *road*, *close*, *mombasa* or *friday*. Other terms such as *kidero* and *grass* make reference to a related controversy created by the governor's decision to plant grass along the roads only a few days before President Obama's visit in an effort to give the city a facelift. The #kiderograss hashtag [4] was used to criticize the situation, as shown in top tweets, that comment with irony about how grass grew overnight, and how this seriously affected traffic. Topic 3 is an example of an ephemeral topic, capturing the spike of reports generated during the visit.

Topic 4 – Accidents ($P(z) = 0.11$). As indicated by the top word (i.e. *accid*), this topic concentrates on tweets related to road accidents. Tweets report the location of the accident (*bypass*), the probable causes for accidents (*caus*, *hit*), the parties involved (*lorri*, *truck*, *car*, *bus*), and the resultant injuries or deaths (*dead*, *injury*).

Topic 5 – Safety Advise ($P(z) = 0.11$). This topic includes comments on safety and tips to fellow commuters, including conversations partly initiated by Ma3route together with partners to inform citizens about various initiatives such as the ZushaLeo initiative on road safety [16] which recommends drivers to stay vigiliant, avoid drinking alcohol while driving, and respect pedestrians (and children in particular) as they cross the street.

Topics are not mutually exclusive, e.g. Topic 1, Topic 2, and Topic 5 capture related issues. For example, there is urban development

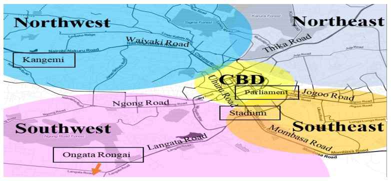

Figure 2: Map with regions discovered in Topics 6-10.

literature that has documented a poor relationship between police and matatu drivers, leading to harassment of matatu operators, and matatu drivers paying bribes to police to avoid harassment [25]. This interplay between Topic 1 and Topic 2 is illustrated in reports that describe checkpoints in the city where police receive bribes from matatu drivers and criticize the consequences of such behavior. In addition, matatus are often criticized for being insecure, which explains why Topic 2 and 5 may overlap.

The remaining topics capture reports about traffic flow. LDA splits reports based on the geolocation of landmarks and type of traffic (slow, moderate, clear, traffic jam, etc.). Figure 2 shows the map of Nairobi with the regions highlighted as in Topics 6-10.

Topic 6 – West Traffic ($P(z) = 0.10$): This topic includes reports of traffic on the west of the city, in the area that extends between Ngong road and Langata road and towards the Ongate Rongai, an informal urban settlement that lies south of Nairobi.

Topic 7 – Southeast Traffic ($P(z) = 0.09$): This topic centers around Jogoo road, Mombasa road, extend including T-mall (south CBD), Mbagahti way and City stadium.

Topic 8 – Northeast Traffic ($P(z) = 0.08$): This topic covers traffic updates and jams in areas surrounding Thika road, Kiambu road, Pangani area and Multhaiga area, and Juja road

Topic 9 – CBD Traffic ($P(z) = 0.08$): This topic covers traffic updates on the central business district encompassing Uhuru highway, Haile Selaissie and Moi avenue, covering important landmarks such as the City Hall and the Parliament. It includes other derivative roads such as Bunyala, Lusaka, and Valley road, in addition to Mombasa road (continuation of Uhuru highway in the south.)

Topic 10 – NorthWest Traffic ($P(z) = 0.08$): This topic includes areas that expands Wayaki way before it converges into Uhuru highway, towards Kangemi (an informal urban settlement in the outskirts of Nairobi) and Limuru (a satellite town north of Nairobi). It also includes Gichuru road, Westland residential district and Chiromo Campus (from Nairobi university), and the museum.

3.4.3 Topics over time. We also found that topics are not only different on their description based on relevant terms but also on their temporal patterns, as observed from their hourly evolution. Figure 3 shows the aggregate hourly distribution for each topic. All topics peak in the morning and evening rush hours. However, we see that Topics 1 to 5 have less abrupt peaks (more regularly spread from 5AM to 7PM), as opposed to traffic-focused topics (6–10), which show stronger decays in the middle of the day. Some topics like topic 6 peak equally in mornings and evenings, while others, like topics 8 and 9 peak more in the morning and the evening hours, respectively. This suggests different temporal traffic patterns.

3.4.4 Topic validation. As an alternative to collecting human annotations to validate the quality of the discovered topics [21], we leveraged the Ma3route categories to compare topics and categories in terms of how many documents are assigned to each topic-category pair. We expect that a well-fitted model would directly capture some of the Ma3route categories, while at the same time we expect broad categories such as *General Info* to be distributed over multiple topics. We focus on the most frequent categories: *General Info*, *Clear*, and *Accident*, and aggregate other ones into one single *Traffic* category. Figure 4 summarizes the tweet assignments for each topic-category combination. First, most posts categorised as *Accident* are indeed captured by our topic 4. This is mainly due to the keywords used to describe these type of events, as opposed to other traffic info. Second, the *General Info* category is distributed over several topics (1, 2, 3, 5 and 6), which indicates that a refinement of such broad category can be obtained. Finally, *Traffic* and *Clear* reports are mainly included in our traffic-focused topics (topics 6 to 10).

In this section, we showed that urban phenomena of interest for multimedia researchers can indeed be studied from social media channels in Global South Cities, and that this blind side could be alleviated if other researchers turned their attention to these regions as sources of relevant research questions. We now move to show the second blind side, this time propagated to machines.

4 BLIND SIDE #2: DEEP VISUAL ANALYSIS

We performed a deep learning-based analysis to explore the kind of visual content shared in the platform. To extract deep learning features from these images, we chose a pre-trained convolutional neural network (CNN) model which uses the GoogLeNet CNN architecture [45] trained on *ImageNet*. To extract the CNN descriptors for each image, we obtained the final layer class probabilities across all 1,000 ImageNet classes. We begin our analysis by examining the distribution of the most likely ImageNet category assigned to each image. Given the probability distribution across categories, we chose the scene category with the highest probability as the dominant class for each image. We performed the analysis on 9,805 images from the 14K Twitter photos (Table 2). Figure 5 shows the histogram of the top-10 recognized ImageNet categories. Overall, the dominant class distribution exhibits long-tail characteristics, with a total of 544 unique ImageNet categories. Around 40% of the images were assigned to these top-10 scene categories.

As illustrated in Figure 5, the dominant category was *minivan* with 8% of the images being classified in this category. Most of the top ten dominant categories are associated with medium to heavy-sized vehicles (*minivan, taxi, trailer trucks, tow trucks, garbage trucks*), which is not surprising given the nature of the traffic content. However, many of the images are misclassified, in particular images classified as *racing car* or *limousine*. Manual inspection reveals that these are not race cars or fancy vehicles. More specifically, images classified as *racing car* correspond to images of cars circulating on dirt roads, cars changing tires on busy roads, and car crashes. Furthermore, images classified as *limousine* correspond to views taken inside cars or matatus that capture the top of the dashboard of these vehicles. These misclassificatons are the result of limited representations of the world encoded in the pretrained CNN models using ImageNet. It has been recently shown that 53%

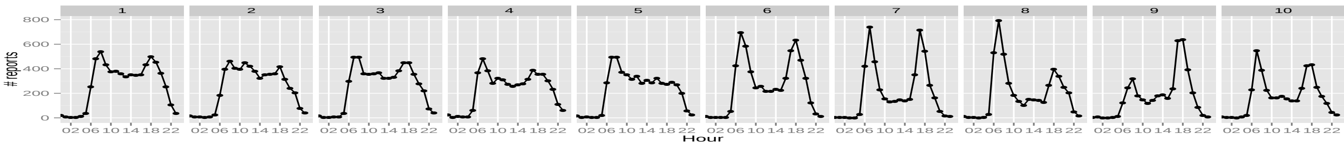

Figure 3: Hourly aggregate of tweets per topic for each one of the 10 topics.

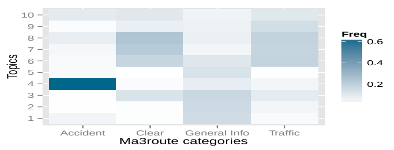

Figure 4: LDA topics vs. Ma3Route categories.

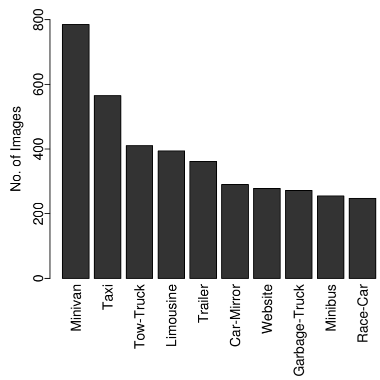

Figure 5: Histogram of Top-10 recognized classes for Ma3Route images that are obtained from Twitter.

of ImageNet images come from only two countries (US and UK) [41]. It seems evident that such sample bias towards wealthy world regions are propagated in our results. It is also troubling that these results mimic results reported in the fairness, accountability and transparency research community about how sample biases affect the recognition of individuals belonging to minority groups [20], but expanding the concerns to urban areas in less wealthy countries. While our analysis will continue in the future, this result demonstrates that pretrained CNNs can reveal basic visual content, but also have limitations with respect to using them to examine images from Global South cities given the sample biases towards wealthy regions. This represents the second blind side, given the current state of multimedia research.

5 AN AGENDA FOR FUTURE WORK

The issues that give rise to the double blind side situation discussed here are complex and multifaceted: while some are technological, others go beyond this realm. We conclude by proposing an agenda for future work that includes tackling the following issues:

The socio-economic bias issue. Although the adoption of mobile phones and mobile internet in Global South cities is increasing, this is a reality that varies across countries, from low-income to lower-middle-income to upper-middle-income countries, using the classification system by the World Bank. Twitter in particular and social media in general are restricted to people having internet access, and therefore our research cannot document the transit issues that affect all citizens. Methods to account for this bias have to be developed, avoiding overgeneralizations.

The data volume issue. Compared to large developed cities, cities in the Global South range from cases where data is produced at similar rates (Istanbul, Sao Paulo, or Mexico City), to cities where data volumes are small, especially for the multimedia research community, accustomed to ever larger data volumes. Africa in particular still generates modest volumes of social media data, but this should not preclude the study of the practices of the current users, as done in this paper. On one hand, we showed that the multimedia data already produced provides rich input to analyze urban dynamic phenomena. Moreover, traffic and public transportation problems are common to many Global South cities, facing challenges related to unsafe driving practices, dangerous road conditions, and overcrowded public transportation. Research could explore transfer learning techniques e.g. to learn data collected from one city and applying it to others of similar characteristics.

The machine perception issue. We showed how pretrained CNNs reveal basic attributes of the image content. While this is already useful, a fundamental question has to do with the need to collect and curate multimedia datasets that reflect urban life in Global South countries, and that can be used to learn models that are better tuned to their realities and alleviate biases, in the spirit of the increasing body of research highlighting the issues of diversity and fairness [48]. We envision the creation of the Global South equivalent of the ImageNet and Places datasets. This would immensely help to make progress on machine perception tasks that are culturally sensitive and useful for these cities.

The human-centered, diversity-aware design issue. We need to improve our understanding of how crowdsourced information is used by citizens in the Global South. This requires a mix of quantitative and qualitative research, and complementary data sources. Do citizens change their transit behaviour or their attitudes based on the information made available in these systems? How much trust is there in the authenticity of systems, specially from the perspectives of fairness and accountability as algorithms become more complex? Finally, can we design tools that could amplify the positive impact of a citizen-driven platform to the community? Answering these questions would require qualitative work interviewing citizens and other urban stakeholders to understand their expectations, beliefs, and usage patterns, and integrating this with multimedia analytics. Human-centered and diversity-aware multimedia could find a wealth of relevant research problems here.

Acknowledgments. This work was supported by the SNSF through the *Dusk2Dawn* project, and by the EC through the H2020 *WeNet* project. We thank the Ma3Route Twitter community for their inspiring civic engagement.

REFERENCES

[1] 2012. The Reluctant Outlaw. http://bit.ly/1OvTV9j.
[2] 2014. Ma3Route Not Moved by Waze 'Launch' in Kenya. http://bit.ly/1LSl93X.
[3] 2015. Bos:311. http://www.cityofboston.gov/DoIT/apps/311.asp.
[4] 2015. Kenyan governor Kidero denies grass removed after Obama. http://www.bbc.com/news/world-africa-33719845.
[5] 2015. Loud music, rude touts make ride in matatus trying for passengers. http://bit.ly/1BBaRlf.
[6] 2015. Major roads to be closed during President Obama's visit. http://bit.ly/1VaGPTj.
[7] 2015. Why Safaricom chose to invest in the waze app. http://www.hapakenya.com/2015/08/14/why-safaricom-chose-to-invest-in-the-waze-app/.
[8] 2016. Why traffic costs Kenya Sh2bn a year. https://www.nation.co.ke/oped/opinion/-Why-traffic-costs-Kenya-Sh2bn-a-year/440808-3183654-format-xhtml-6xsy0bz/index.html.
[9] 2019. FixMyStreet. https://www.fixmystreet.com.
[10] 2019. ma2too3a. http://ma2too3a.com/.
[11] 2019. Ma3Route. http://www.ma3route.com/.
[12] 2019. SeeClickFix. http://seeclickfix.com.
[13] 2019. TrafficDito. http://www.trafficdito.com/.
[14] 2019. Traffline. http://www.traffline.com/.
[15] 2019. Waze. https://www.waze.com/.
[16] 2019. Zusha. https://twitter.com/zusharoadsafety.
[17] Yemisi Akinbobola. 2012. Social media stimulates Nigerian debate on sexual violence. http://bit.ly/1eONoA2.
[18] Michael L Best and Amanda Meng. 2015. Twitter democracy: policy versus identity politics in three emerging African democracies. In *Proc. of Inter. Conf. of Information and Communication Technologies for Disaster Management*. ACM.
[19] D.M. Blei. 2012. Probabilistic topic models. *Commun. ACM* 55, 4 (2012), 77–84.
[20] Joy Buolamwini and Timnit Gebru. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. In *Proc. ACM Conf. on Fairness, Accountability and Transparency (FAT)*.
[21] Jonathan Chang, Sean Gerrish, Chong Wang, Jordan L Boyd-Graber, and David M Blei. 2009. Reading tea leaves: How humans interpret topic models. In *Advances in neural information processing systems*. 288–296.
[22] Marco De Nadai et al. 2016. Are Safer Looking Neighborhoods More Lively?: A Multimodal Investigation into Urban Life. In *Proc. ACM MM*.
[23] Carl Doersch, Saurabh Singh, Abhinav Gupta, Josef Sivic, and Alexei A. Efros. 2012. What Makes Paris Look Like Paris? *ACM Trans. Graph.* (2012).
[24] Huiji Gao et al. 2011. Harnessing the crowdsourcing power of social media for disaster relief. *IEEE Intelligent Systems* (2011).
[25] Jennifer Graeff. 2009. The organization and future of the matatu industry in Nairobi, Kenya. *Paper prepared by Center for Sustainable Urban development, earth Institute, Columbia University, New York* (2009).
[26] Thomas L Griffiths and Mark Steyvers. 2004. Finding scientific topics. *Proc. of the National Academy of Sciences* 101 (2004), 5228–5235.
[27] Kurt Hornik and Bettina Grün. 2011. topicmodels: An R package for fitting topic models. *Journal of Statistical Software* 40, 13 (2011), 1–30.
[28] Philip N Howard, Aiden Duffy, Deen Freelon, Muzammil M Hussain, Will Mari, and Marwa Mazaid. 2011. Opening closed regimes: what was the role of social media during the Arab Spring? *Available at SSRN 2595096* (2011).
[29] Kokil Jaidka and Saifuddin Ahmed. 2015. The 2014 Indian General Election on Twitter: An Analysis of Changing Political Traditions. In *Proc. ICTD*.
[30] Lyndon Kennedy et al. 2007. How Flickr Helps Us Make Sense of the World: Context and Content in Community-contributed Media Collections. In *Proc. ACM MM*.
[31] Neha Kumar. 2014. Facebook for self-empowerment? A study of Facebook adoption in urban India. *New media & society* (2014).
[32] Andres et al. Monroy-Hernandez. 2013. The new war correspondents: The rise of civic media curation in urban warfare. In *Proc. of CSCW*.
[33] Kenda Mutongi. 2006. Thugs or entrepreneurs? Perceptions of matatu operators in Nairobi, 1970 to the present. *Africa* 76, 04 (2006), 549–568.
[34] Mor Naaman, Amy Zhang, Samuel Brody, and Gilad Lotan. 2012. On the Study of Diurnal Urban Routines on Twitter.
[35] Vicente Ordonez and Tamara L. Berg. 2014. Learning High-Level Judgments of Urban Perception. In *Proc. ECCV*.
[36] Lorenzo Porzi et al. 2015. Predicting and understanding urban perception with convolutional neural networks. In *Proc. ACM MM*.
[37] Daniele Quercia, Neil Keith O'Hare, and Henriette Cramer. 2014. Aesthetic Capital: What Makes London Look Beautiful, Quiet, and Happy?. In *Proc. CSCW*.
[38] Daniel Ramage, Susan T Dumais, and Daniel J Liebling. 2010. Characterizing Microblogs with Topic Models. *ICWSM* (2010).
[39] Salvador Ruiz-Correa, Darshan Santani, Beatriz Ramirez-Salazar, Itzia Ruiz-Correa, Fatima Rendon-Huerta, Carlo Olmos Carrillo, Brisa Sandoval Mexicano, Angel Arcos Garcia, Rogelio Hasimoto-Beltran, and Daniel Gatica-Perez. 2017. SenseCityVity: Mobile Crowdsourcing, Urban Awareness, and Collective Action in Mexico. *IEEE Pervasive Computingg, Special Issue on Smart Cities* 16, 2 (Between April and June 2017), 44–53.
[40] Darshan Santani, Jidraph Njuguna, Tierra Bills, Aisha W. Bryant, Reginald Bryant, Jonathan Ledgard, and Daniel Gatica-Perez. 2015. CommuniSense: Crowdsourcing Road Hazards in Nairobi. In *Proc. MobileHCI*.
[41] S. Shankar, Y. Halpern, E. Breck, J. Atwood, J. Wilson, and D. Sculley. 2017. No Classification without Representation: Assessing Geodiversity Issues in Open Data Sets for the Developing World. In *Proc. NIPS Workshop on Machine Learning for the Developing World*.
[42] Carson Sievert and Kenneth E Shirley. 2014. LDAvis: A method for visualizing and interpreting topics. In *Proc. of the Workshop on Interactive Language Learning, Visualization, and Interfaces*. 63–70.
[43] Tomer Simon, Avishay Goldberg, Limor Aharonson-Daniel, Dmitry Leykin, and Bruria Adini. 2014. Twitter in the Cross Fire—The Use of Social Media in the Westgate Mall Terror Attack in Kenya. *PLoS ONE* (2014).
[44] Fredrik Matias Sjoberg, Jonathan Mellon, and Tiago Peixoto. 2015. The Effect of Government Responsiveness on Future Political Participation. *Available at SSRN 2570898* (2015).
[45] Christian Szegedy et al. 2015. Going deeper with convolutions. In *Proc. CVPR*.
[46] L. Townsend and C. Wallace. 2016. Social Media Research: A Guide to Ethics. *University of Aberdeen* (2016).
[47] Kentaro Toyama. 2015. *Geek heresy: rescuing social change from the cult of technology*. Public Affairs.
[48] S.M. West, M. Whittaker, and K. Crawford. 2019. Discriminating Systems: Gender, Race and Power in AI. *AI Now Institute* (2019).
[49] Wikipedia. 2019. List of urban areas by population. https://en.wikipedia.org/wiki/List_of_urban_areas_by_population.
[50] Sarah Williams, Adam White, Peter Waiganjo, Daniel Orwa, and Jacqueline Klopp. 2015. The digital matatu project: Using cell phones to create an open source data for Nairobi's semi-formal bus system. *Journal of Transport Geography* 49 (2015), 39 – 51.
[51] Susan P Wyche, Andrea Forte, and Sarita Yardi Schoenebeck. 2013. Hustling online: understanding consolidated facebook use in an informal settlement in Nairobi. In *Proc. CHI*.
[52] Dave Yates and Scott Paquette. 2011. Emergency knowledge management and social media technologies: A case study of the 2010 Haitian earthquake. *International journal of information management* (2011).
[53] Wayne Xin Zhao, Jing Jiang, Jianshu Weng, Jing He, Ee-Peng Lim, Hongfei Yan, and Xiaoming Li. 2011. Comparing twitter and traditional media using topic models. In *Advances in Information Retrieval*. Springer.