Chapter 12
Leveraging Cellphones for Wayfinding and Journey Planning in Semi-formal Bus Systems: Lessons from Digital Matatus in Nairobi

Jacqueline Klopp, Sarah Williams, Peter Waiganjo, Daniel Orwa and Adam White

Abstract For many cities in the developing world, public transit consists mainly of semi-formal mini-buses (paratransit). However, little to no digital information is typically available on routes, bus stops, passenger boarding, service frequency or scheduled trip times. Cities that rely on these bus systems can benefit from the generation of digital data on these systems for planning and passenger information purposes. Perhaps more importantly, this data can provide the ability to generate citizen-based information tools, such as transit routing applications for mobile devices widely discussed in the smart cities dialog (Townsend in Re-Programming mobility: the digital transformation of transportation in the United States, 2014). Through our work in Nairobi, this paper shows that cell-phone technology that is ubiquitous in most countries can be used to generate a dataset in an open standard, GTFS. Citizens can then leverage open source tools made for that standard, enhancing access to information about the transit system. We argue that one of the

J. Klopp
Center for Sustainable Development, Columbia University, 475 Riverside Dr. Suite 520, New York, NY 10115, USA
e-mail: jk2002@columbia.edu

S. Williams (✉)
Department of Urban Studies and Planning, Massachusetts Institute of Technology, 77 Massachustts Ave, Cambridge, MA 02139, USA
e-mail: sew@mit.edu

P. Waiganjo · D. Orwa
School of Computing and Informatics, University of Nairobi, 30197 Nairobi GPO 00100, Kenya
e-mail: waiganjo@uonbi.ac.ke

D. Orwa
e-mail: dorwa@uonbi.ac.ke

A. White
Groupshot, 157 Pearl Street, Cambridge, MA 02139, USA
e-mail: adam@groupshot.org

S. Geertman et al. (eds.), *Planning Support Systems and Smart Cities*, Lecture Notes in Geoinformation and Cartography,
DOI 10.1007/978-3-319-18368-8_12

most important components of our work in Nairobi was the engagement process that created trust in the data and knowledge of its existence for the development of civic technologies. The lessons learned in Nairobi can be translated to other areas with the potential to use mobile applications to develop data on essential urban infrastructure and to extend the use of that data by sharing it with a larger community.

1 Introduction

In this chapter, we describe a research effort in Nairobi, Kenya to collect basic data on the semi-formal bus system that is the basis of the public transit system in this city of over 3.5 million people. These buses of all sizes are called "matatus", and they are ubiquitous in Nairobi, forming the bulk of the mass transit system. We show that data collection using GPS-enabled cell phone applications can generate valuable data that allows for the development of wayfinding tools, both digital and analog, for transit users in the matatu system. These tools include trip planning applications on mobile devices and a stylized paper map of bus routes produced from the data. Our research illustrates how by actively engaging the Nairobi technology and transit community during the data collection process we were able to extend the benefits of the data to multiple actors in the city, thereby extending the benefit of the data collection process itself.

Increasingly, formal transit authorities are releasing data to facilitate the development of tools to help transit users make more informed choices, which, in turn, can improve transit systems as a whole (Roth 2010; McHugh 2013). Access to data is part of a broader agenda of leveraging technology and analytics to make cities including their transit systems "smart" (Townsend 2013). However, in many rapidly developing cities where people rely on semi-formal bus systems operated by diverse actors, access to information about these systems is not readily available. Often regulation and planning is poor which means operators are not required to collect information similar to more formal transit agencies. Even if these operators were required to collect data, the task of gathering this data from many small operators would be challenging. Each operator would have to collect data in the same standard and would have to ensure the accuracy and quality control (Williams et al. forthcoming).

Lack of data has made it hard to include semi-formal transit systems in transportation models. However, these systems often make up the majority of the vehicles on the roadway and, therefore, are essential for accurately representing traffic patterns in these cities. The limited access to information also makes it hard for users to know how to navigate the systems, and they often must rely on word of mouth for information about stops, routes, and schedules.

Given that data is essential for planning and citizen knowledge of semi-formal transit systems, it is important to develop technologies that allow for ease in data collection. The rapid growth of mobile phone use along with the increased accuracy of GPS transceivers in mobile devices allow them to be more easily used for data collection (Hein et al. 2008). A number of recent experiments have demonstrated that using cell phone transaction logs, which include location information, can assist in mapping traffic flows (Caceres et al. 2012; Herrara et al. 2010; Wang 2010; Ratti et al. 2006) and empower transit users in many ways "exemplifying the smart city vision" (Zegras et al. 2014). However, telecommunications companies, who collect these records are often unwilling to release cell phone data, even when anonymized, citing privacy concerns. One recent exception is the analysis of data from 2.5 billion call records of 5 million cell phone users in the Ivory Coast which helped analyze and suggest improvements to traffic movement in and around Abidjan, Ivory Coast (Talbot 2013; Wakefield 2013).

While access to cell phone use records can be difficult, the information can also be collected through direct user interaction with cell phones. In our research in Nairobi, we developed cell phone based mobile applications that made it easier to collect location data on routes, trips, fares, stops and schedules using a small group of dedicated people. After we had begun our work in Nairobi, we found several other projects that have attempted the same process, in Dhaka, Manila, and Mexico City (Eros et al. 2014; Klopp et al. 2014; Zegras et al. 2014). However, with the exception of Manila these projects did not open the semi-formal transit data for anyone to use.

In our project in Nairobi, we developed and released the data on the semi-formal transit system in an open transit standard, GTFS. Openly sharing data in this way allows it to be interpreted by multiple groups and extends the knowledge production that can be created from it (Craig 2005; Williams et al. 2014). The data collection process included continuous communications with the local technology and transit community in Nairobi. Through this process, the technology community learned about GTFS and that it could be used with open source tools for trip routing, such as Open Trip Planner among others. GTFS is the standard used for most routing applications. It was originally developed by Google and Portland's Tri-met transit agency to allow for transit routing in Google maps (McHugh 2013). Beyond routing applications, there are also a number of open source analysis packages that use GTFS as the base (Wong 2013). We distributed the data on the GTFS exchange, a website repository for open transit data, as well as our website www.digitalmatatus.com, allowing anyone to access it and develop projects using it.

By informing Nairobi's transit community about the development of the GTFS data throughout the data collection process, we were able to build trust in the data set created. Public Participation GIS (PPGIS) scholarship has shown this trust is essential for the adoption and use of data (Elwood 2006). We held a series of workshops where we invited actors from the local technology community, governments, NGOs, academia, as well as operators and drivers of the semi-formal buses. At these workshops, we explained the data collection process. Participants also learned how they could access and use the information. By developing this inclusive process, we broke down the knowledge barriers often established by the control and dissemination of information (Harvey and Christman 1998; Sieber 2006; Williams et al. 2014).

Mobile technology can serve as a powerful tool for data collection and feedback for next generation PPGIS projects, given that processes for collecting data can create new forms of knowledge production (Elwood et al. 2012). Smartphones now come with several sensors including GPS receivers; this means that smartphone users can contribute GPS tagged data with relative ease, which, can in turn, be used to develop data beneficial for everyone. Data collection on mobile devices can be exploited using a controlled collection process to assure quality data for a multiplicity of actors and users, especially if it is made open (Ching 2013; Klopp et al. 2014; Williams et al. 2014; Zegras et al. 2014). The cellphone can thus be thought of as the "quintessential smart city instrument" (Zegras 2014) although, as always, how this instrument will be used, just like PPGIS as a whole, is heavily linked to context and politics (Goodspeed 2015; Ghose and Elwood 2006; Kyem 2004).

In the last decade, citizen participation in the development of crowdsourced data collection tools has shown the power of data collection that can often be performed outside formal governmental channels. The push for open data, such as GTFS for transit, as well as open software means that the private sector can more easily be involved in the development of smart technologies that can enhance citizen interaction with the city (Townsend 2013). This is particularly true for the development of mobile phone technologies which can democratize access to essential information about city resources.

2 Methodology

Our team, consisting of a consortium of universities (University of Nairobi, MIT and Columbia University) and one small technology firm (*Groupshot*), mapped the city's bus (matatu) routes, primarily using easily available smartphone apps. Data collection took place by having students directly riding a matatu route, either in the car or matatu, using a data collection tool to generate latitude and longitude points along the route. A record of all of the stops as well as specific metadata about each stop, such as the stop name, stop signage and stop infrastructure (shelter and bench) was also recorded. While some semi-formal bus systems do not utilize fixed stops, in Nairobi the matatu system features regular stops and large terminals. We identified a stop based on local knowledge of data collectors, information from frequent users of these routes, visual notation (signs, shelters etc.), and if necessary, confirmation from discussion with an operator or group of commuters on a route.

Prior to beginning of data collection, we surveyed existing bus data and found that no complete up to date publicly accessible data existed. The city had MS Word document files for some routes dated 1997 and some app developers had tried to collect data using GPS devices but this was for individual routes and did not represent a comprehensive data set. Other groups like the weekly magazine of events in Nairobi, Kenya *Buzz*, created a stylized matatu map from scratch but it had many geographical inaccuracies and was not based on latitude and longitude points, making it hard for the information to be used with digital technologies. However, these attempts showed demand for this data from different constituencies.

In the fall of 2012, we tested various software developed for data collection on cell phones to determine one suitable for collecting route and stop data of Nairobi's matatu system (see Table 1). We compared the latitude and longitude data collected on these systems to data collected by hand-held GPS units to determine accuracy. Many of the devices captured data comparable to GPS units. However, some lost connection with the satellites which created incomplete data. On the basis of this review, the team decided to use the *MyTracks* App developed by Google for Android phones. *MyTracks* was chosen because it allows users to enter the name of stops digitally. It also snaps latitude and longitude points to the Google base dataset that is used as the map interface for the program. The ability to snap the GPS data to roads allowed for a cleaner data set as the GPS receivers often have errors of 10 m. Snapping to the roads allowed us to ensure our data on stops and routes was in the actual roadway making for a much cleaner dataset. The *MyTracks* application also made it easy to add our own metadata to the data, including information about whether stops were designated or undesignated.

TransitWand, an open source mobile data collection software for Android phones developed by the firm Conveyal with support from the World Bank, was released in the last weeks of our data collection and had features similar to *MyTracks* in that it allows for more metadata collection and the ability to snap GPS points to the roadway. The team investigated the use of *TransitWand* for the project. However, the tool was not suitable for our collection process as it needed modification for use in Nairobi. These modifications included: developing a script that would allow the program to generate unique ids using the structure we had created specifically for our Nairobi GTFS data, extending the code to create data columns and input forms to accommodate this way we extended the GTFS data format for Nairobi, and modifying the underlying base interaction as in our preliminary test the latitude and longitude points were not snapping to the roads properly. These modifications were not extensive. However given the fact that we were at the end of our data collection process and the export to GTFS function was not working properly in *TransitWand*, we did not shift to using this data collection device. Future research should test the modification and adaptation of this tool, because its interface and expandable metadata collection in the field holds much potential.

Formal data collection began in February of 2013 using a modified version of *MyTracks* App for Android smart phones developed specifically for the project to allow for the entry of stop names, as well as identifying if a stop was designated or undesignated. (see Fig. 1) GPS units were used as backup so that we could compare the accuracy of the two forms of data collection. Through testing we found that GPS units and data collected through *MyTracks* had similar accuracy. Some of the challenges with using this mobile based application was the limited battery life, slow speeds of affordable Android phones, the risk of losing more high value Android phones to theft in matatus where security could be a challenge (that, did happen, unfortunately), and the small screen size that made digital data entry more time consuming—particularly in situations where bus stops were very close to one

Table 1 Comparison of various mobile collection software tested for project

Tool tested	Pro	Con
MyTracks	• Snapped latitude and longitude data to the roadways allowing for cleaner data.	• Needed to develop tool to post process data collected in the field to GTFS data standard.
	• Easy to modify for additional data elements.	
	• Easy to visualize the data captured immediately.	
TransitWand	• Snapped latitude and longitude data to the roadways allowing for cleaner data.	• Needed programming modifications for use in the field which was already developed for MyTracks. However had this tool been found earlier in the process these modifications might have been very similar to what was needed to be performed by MyTracks.
	• Developed specifically for the collection of GTFS data.	• GTFS export was not working properly making post processing to GTFS necessary and cumbersome.
	• Allowed for modification of questions asked in the field.	
	• Good user interface.	
FlockTracker	• Allowed for latitude and longitude data collection on phones.	• Needed many modifications to be used in Nairobi. A new interface would need to be developed.
	• Could be modified to allow for survey questions.	• As this tool was in the early phase of development when we used it, programming communication errors causes it to lose connection with GPS satellites. We understand some of these bugs have been fixed in more recent versions of the software.
		• Did not snap data to the roadways - which meant that much post-processing would be needed.
Fulcrum	• Allowed for latitude and longitude data collection on phones.	• Designed for development on iPhones which are not common in Kenya as they are too expensive. Most smartphone users in Kenya use Android phones.
	• Could be modified to allow for survey questions.	
GPS surveyor	• Allowed for latitude and longitude data collection on phones.	• Works much like a GPS unit and does not allow for additional data (such as stop names to be collected in the field) no advantage over GPS units.
Open data kit	• Allowed for latitude and longitude data collection on phones.	• Needed significant programming development to allow for use in the field.
App inventor		• Needed significant programming development to allow for use in the

(continued)

Table 1 (continued)

Tool tested	Pro	Con
	• Allowed for latitude and longitude data collection on phones.	field. Was the basis for the FlockTraker program previously mentioned.

It should be noted that for all the products we tested, the interaction between the phone and GPS satellites needed for data collection drained the battery life of the phones, making it difficult to complete whole routes without extra battery packs for the phones themselves

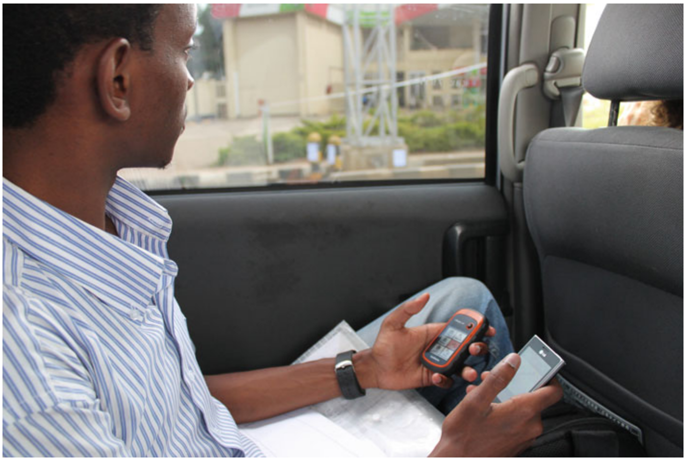

Fig. 1 University of Nairobi Student collecting GTFS data using *MyTracks* software and GPS unit. Image credit Adam White

another. A clear opportunity exists to continue to explore, compare and improve this transit data collection technology.

Once the data was collected, it needed to be cleaned and formatted into the standard we had chosen-the General Transit Feed Specification (GTFS) used by Google (see Fig. 2). The GTFS standard was chosen as it is the most common form of openly shared transit data. There are several open source applications, from routing to accessibility analysis, that use this format. Levering the existing technology developed for this format would extend its use in Nairobi. Translating the data into GTFS included the development of a unique identification system for routes and stops. There were also a number of challenges including the fact that GTFS required scheduling information, which matatus loosely follow, so a hypothetical schedule had to be generated in order for the data to be used with routing software. We also needed to develop new fields to include information we collected about designated and undesignated routes (Williams et al. forthcoming).

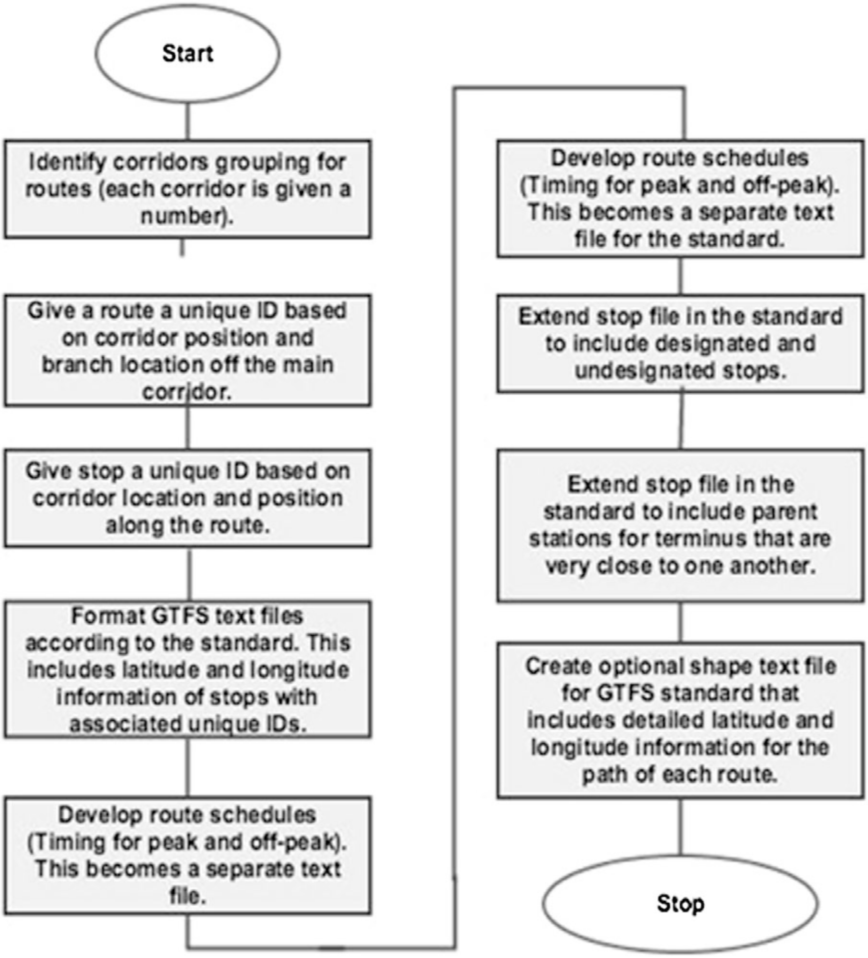

Fig. 2 Flowchart for the conversation of data collected into GTFS format

As we collected this data, we held a series of conversations with potential users in the technology and transit community to help them become aware of the data, its structure and format, and how they might be able to use it. We also used these discussions to get feedback on our process. In all, we collected 136 routes and 6 rail lines and made this data open on our website (www.digitalmatatus.com), Facebook page, and the GTFS Exchange. Making this data open was part of our strategy to encourage technology to be developed on top of the data and hence to generate innovation in wayfinding, journey planning for users, and more analytical applications for the transportation community.

In order to encourage the use and adoption of the data to users inside and outside the transit community, it was necessary to show community members what was

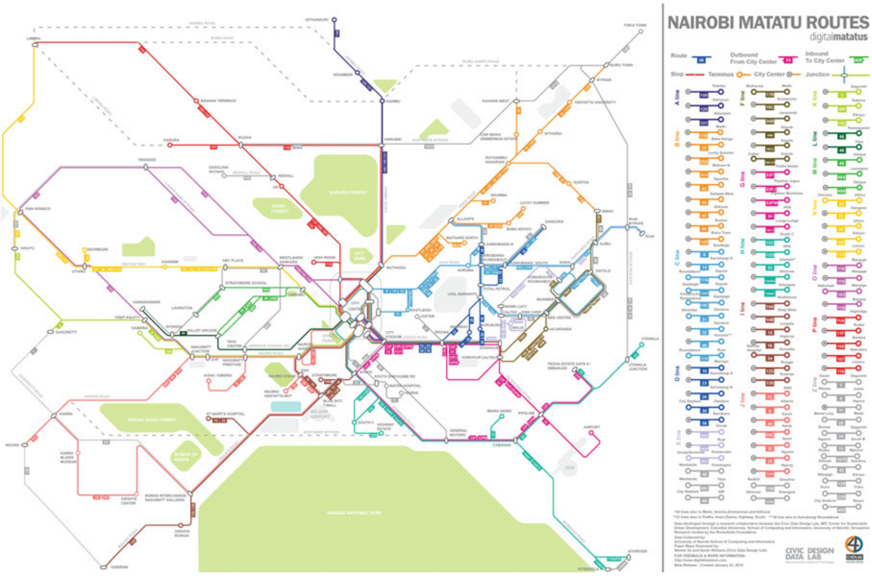

Fig. 3 Stylized Nairobi Matatu route map created through the project

Fig. 4 Image of team during an editing session. Image Credit: Adam White

collected through visualization of the data. We thus created a paper version of the matatu data as a way to orient the public. It borrowed its style from traditional transit maps such as those found in London, Paris, New York, and Washington DC (Fig. 3). University of Nairobi students helped us edit the map to correct for errors, identify missing routes and review overall legibility (see Fig. 4). The maps transformed how Nairobi's transit community understood the matatu system, because, for the first time, they could see it as one comprehensive network. Matatu owners used the map to illustrate their strategies for relieving congestion. The launch of the data and map, where the governor of Nairobi made it the official Matatu map, helped start a meaningful dialogue about the role of matatus in Nairobi's transit landscape. The downloadable maps went viral on social media, and versions were printed in local newspapers allowing anyone to access the information we created.

After the release of the data and map, we held two focus groups to solicit feedback on the map's accuracy, usability and value. One focus group was with twenty-five university students who were regular transit users, and another focus group was with sixteen matatu drivers and owners who worked on one of the Southern routes. We also conducted a more formal survey of first-year planning students at the Technical University of Kenya. A total of sixty-six surveys were collected after respondents had the opportunity to review and comment on the paper version of the matatu map. Respondents were asked a total of nineteen questions which included demographic information on participants as well as overall familiarity with matatus.

3 Results

Our Nairobi study illustrates that by modifying standard software developed for Android mobile phones, it is possible to develop a comprehensive data set for semi-formal bus systems. The data collection requires careful groundwork identifying and riding all possible routes. The team also repeated travel in some routes to locate the most common path, as these are prone to some variability. Once developed, the dataset provides base data that can be built upon and updated. We are currently researching and developing methods including crowdsourcing and new tools to easily update changes to routes and stops.

The release of the data showed the demand for the data by planners and researchers working in Nairobi. We received multiple requests for the data from transportation consultants including the Institute for Transportation and Development Policy (ITDP) which was creating a bus service plan for a proposed Bus Rapid Transit (BRT) corridor, and required knowledge of existing routes. Providing the data saved the organization much time as a lot of work went into investigating the various matatu routes along the corridor which was the primary emphasis for their study. They used the data and map to select routes and locations on these routes to conduct ridership capacity and frequency surveys (ITDP and UN-Habitat 2014). The matatu owners and drivers instantly saw the potential of the map and used it to demonstrate the logic of new routes that would help improve traffic circulation. The consultants at Mott McDonald who were advising the Ministry of Transport and Infrastructure in setting up a new transit authority also found the data valuable for their purposes. Numerous researchers used the data and maps for their work. Providing the data openly allowed the transit community to share ideas from a common knowledge base creating more dialogue, improved trust and enhanced collaboration between some of the various stakeholders (Williams et al. 2014).

Our strategy of developing open data for innovation was very successful. A number of software developers including those we included in the early discussions about the project used the data to develop trip planning applications for cell phones. These included the popular trip planning application *Ma3route* which also gives traffic alerts and allows you to rate your matatu driver. *FlashCast Sonar* was another trip planning application that also gave some real time information for a subset of the matatus in Nairobi. The *Digital Matatu* trip planning app was created for Windows, named presumably after this research project and another trip planning app called *Matatu Map* was created. More recently a Canadian company used our data to produce *Transit App*. Research remains to be done on how many users these apps attract and the impacts of their use. *Ma3route* also crowdsources information about traffic conditions, creating potentially useful data for planning. Google queries may also be interesting to analyze once the data is live on their system.

The stylized matatu map was extremely well-received when launched in Nairobi and the city intended to adopt it as its official map. The map was used in a number of transit planning initiatives. Citizens were generally very positive about the map including the 66 undergraduate planning students from the Technical University of Kenya who took our formal survey. The survey sample did not reflect the makeup of the city- a more representative sample would be needed. However, the results are still helpful to gauge the response of a young better educated male demographic. About two-thirds of respondents were under the age of 22 years of age, and nearly 88 % were males. Most respondents (48 %) had lived in Nairobi for between two and ten years and most rode matatus four or more times per week, (42 %). Almost 29 % claimed only to ride matatus a few times per month reflecting the fact that many of the students walk.

More than 80 % of respondents reported never having seen a public transit map like the Digital Matatu map for their city. More than 90 % claimed that they were able to find where they were at the time and also the route they typically rode. Almost all, 86 %, believed the map made it easier to find the way around the city. About 83 % reported that with the map they are more likely to take a new matatu route and 77 % reported that they were more likely to take a matatu to a part of Nairobi they usually did not travel to. In focus group discussions, it was clear that finding information about new areas of the city was a problem with people needing to rely on word of mouth. In addition, respondents agreed that the matatu map would be useful for a variety of users: visitors and tourists (19 %), newcomers to the city (14 %), residents (12 %), and matatu users/passengers (11 %). Nearly 45 % of respondents reported that posting the map at matatu stops would be the most convenient way to access the map, and this corresponded to what we learned in the focus groups with drivers who had a similar request. One respondent suggested posting the map in public places like supermarkets and shops. Overall, the results showed an interest in having public transit information in map form.

Several respondents hoped that real-time information could be incorporated in the data some way, which the team would like to investigate in the future. As the cost of GPS devices lowers further, more vehicles could carry such devices and give real time information to passengers although many vehicles do not wish to be tracked when they deviate from designated routes or make more trips than they might report to the matatu owner. The *FlashCast* sonar app already provides real-time data for a several matatu routes.

It is interesting that almost 50 % of survey respondents claimed that after seeing the map, they believed that Nairobi's matatu system was better than they previously thought. The stylized representation of the map helps give that impression as it copies the style of more formalized systems but also shows the expanse and network structure of the system which was previously hard to understand. About 86 % reported that the Digital Matatu Map was more useful than their available sources of matatu information, the most typical of which for 42 % of respondents was to ask a friend. Nearly two-thirds of respondents claimed that they would be willing to pay between KES 20 (US $0.22) and KES 100 (US $1.10) for the map. While we need to survey a representative sample of the Nairobi population, these initial results of young primarily male university students suggest that a demand exists for better transit information within this demographic category. Further research should explore how other ways of disseminating the map data might provide better information for users and the public to understand and relate to their transit system (Gosselin 2011).

Finally, our focus group with drivers and owners of matatus revealed a willingness to participate in improving data on landmarks and stops on the map. Like the university respondents, they thought the map was generally useful and wanted to have parts of it at bus terminals and certain route information available on their vehicles so people knew the stops. They said that they spent time giving directions to passengers, and they did not always have good information on routes in other parts of the city. They took a certain pride in creating and running part of the extensive system depicted on the map.

4 Conclusions

Overall, this work in Nairobi gives preliminary evidence that specialized data recording phone apps, utilizing the integrated GPS tracking function of mobile phones, can effectively and efficiently capture important transit information. Our research so far also demonstrates demand exists for both the open data and information in both digital and paper forms. More importantly, we illustrate the importance of engaging the transportation and technology community in data collection strategies. In the era of "big data" and the prevalence of crowdsourcing devices, our project demonstrates how valuable data can be collected and disseminated in a way that extends the possibilities of innovation in civic technologies by creating networks of trust and collaboration.

Our research also suggests that semi-formal bus systems that are fragmented and operated by many vehicles, could benefit from a more centralized data collection effort by a university or transit authority (which currently does not exist in Nairobi but is now being created). This would allow for data quality control and the development of standard such as GTFS. We have also shown that universities in these cities can help with these efforts and can help spur innovation by linking to the tech community. However, substantial challenges remain in institutionalizing data collection and supporting better official city efforts to provide needed transit information to citizens in varied forms that cater to different categories of people including the poor, the elderly and children. Currently, such an effort does not exist because officials in Nairobi still have negative attitudes to their "semi-formal" transit system (Klopp and Mitullah 2015), and generally do not have a culture of providing information to citizens (Klopp et al. 2013; Williams et al. 2014) although this is changing.

In the movement to develop cities that are "smarter" and more responsive to citizens' needs, the development of data on essential infrastructure such as the matatus is critical. Better wayfinding and journey planning along with new planning applications can help Nairobi's transit system become more responsive to citizen needs. Yet there is much work to do to link this information up with transportation planning processes. Wayfinding, for example, is highly dependent not only on technology and information but on the structure of the environment and the transportation network (Woyciechowicz and Shliselberg 2005). However, we have shown that cell phone technology is making it easier to help users and operators even in these highly flexible and dynamic semi-formal bus systems. The kinds of digital data we collected, the openness in which it was released and shared, and the visualization of the network of routes are important steps but not the only ones necessary for better transit information.

5 Recommendations

Overall, this work suggests the fruitfulness of transit data collection efforts using GPS enabled cell phone devices. However, more research needs to be done to improve transit data applications, data collection, visualization, and dissemination efforts to make transit information widely available in appropriate forms. This in turn, will be critical for improving wayfinding and journey planning but also analysis and modeling to develop interventions to improve existing transit systems. More effective phone applications for easy, user-friendly data collection might be developed that streamline information processing while in the field. With good quality tools and continued engagement with residents, government, and civic "hacktivists" (Townsend 2013) there are opportunities to crowdsource important data from citizens. More research needs to be performed on what additional transit information is helpful to users, operators, and planners. Further research should also investigate the potential impacts that providing this data and information has on the relationships between the various actors of the transportation network.

Acknowledgments This work would not have been possible with the critical thinking and hard work of the University of Nairobi, School of Computing & Informatics students: Ikamar Ekessa, Peter Kamiri, Samuel Kariu, Maureen Mbinya, Jackson Mutua, Mureri Ntwiga. Researchers at MIT's Civic Data Design Lab also contributed significantly to the work including; Jonathan Andrew Campbell, Emily Eros, Alexis Howland, Lindiwe Rennert, Alicia Rouault, Christopher Van Alstyne, Catherine Vanderwaart. Special thanks to Wenfei Xu from the Civic Data Design lab who was instrumental in the development of the Matatu Map. Thanks also to Professor Lawrence Esho and his students at the Technical University of Kenya for their valuable feedback on the Digital Matatu map. We also gratefully acknowledge the support of the Rockefeller Foundation which provided a grant for this work. Special thanks to Benjamin de la Pena at the Rockefeller Foundation for his support of and critical insights on this work. We also benefitted greatly from conversations that compared projects with Prof. Chris Zegras, Albert Ching, Stephen Kennedy, Neil Taylor, Kevin Webb and Emily Jean Eros. Last but not least, we thank James Gachanja and Dr. Zachary Gariy at KIPPRA for hosting the workshops and the Kenya Alliance of Resident Associations for their work on the launch of the data and map for this work helped our thinking, made connections and showed us a way forward in sustaining data collection work.

References

Caceres, N., Romero, L. M., Benitez, F. G., & Del Castillo, J. M. (2012). Traffic flow estimation models using cellular phone data. *Intelligent Transportation Systems, IEEE Transactions on, 13*(3), 1430–1441.

Ching, A., Zegras, C., Kennedy, S., Mamun, M. (2013). A user-flocksourced bus experiment in Dhaka: New data collection technique with smartphones. *Transportation Research Record: Journal of the Transportation Research Board*.

Craig, W. (2005). White knights of spatial data infrastructure: The role and motivation of key individuals. *URISA Journal, 16*(2), 5–13.

Elwood, S. (2006). Negotiating knowledge production: The everyday inclusions, exclusions, and contradictions of participatory GIS research. *The Professional Geographer, 58*(2), 197–208.

Elwood, S., Goodchild, M. F., & Sui, D. Z. (2012). Researching volunteered geographic information: Spatial data, geographic research, and new social practice. *Annals of the Association of American Geographers, 102*(3), 571–590.

Eros, E., Mehndiratta, S., Zegras, C., Webb, K., Ochoa, M. C. (2014). Applying the general transit feed specification (GTFS) to the global south: Experiences in Mexico City and beyond. *Transportation Research Record*.

Goodspeed, R. (2015). Smart cities: Moving beyond urban cybernetics to tackle wicked problems. *Cambridge Journal of Regions, Economy and Society*, *8*(1), 79–92.

Gosselin, K. (2011). Study finds access to real-time mobile information could raise the status of public transit. *Next American City*. http://nextcity.org/daily/entry/study-finds-access-to-real-time-mobile-information-could-raise-the-status-o. Accessed March 6, 2015.

Harvey, F., & Chrisman, N. (1998). Boundary objects and the social construction of GIS technology. *Environment and Planning A, 30*(9), 1683–1694.

Hein, J. R., Evans, J., & Jones, P. (2008). Mobile methodologies: Theory, technology and practice. *Geography Compass, 2*(5), 1266–1285.

Herrera, J. C., Work, D. B., Herring, R., Ban, X. J., Jacobson, Q., & Bayen, A. M. (2010). Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment. *Transportation Research Part C: Emerging Technologies, 18*(4), 568–583.

Klopp, J. M., Marcello, E. M., Kirui, G., Wing, N., & Mwangi, H. (2013). Can the internet improve local governance? The case of Ruiru, Kenya. *Information Polity, 18*(1), 21–42.

Klopp, J., Mutua, J., Orwa, D., Waiganjo, P., White, A., Williams, S. (2014). Towards a standard for paratransit data: Lessons from developing GTFS data for Nairobi's Matatu system. In *Transportation Research Board 93rd Annual Meeting* (No. 14-5280).

Klopp J., & Mitullah, W. (2015). Politics, policy and paratransit: A view from Nairobi. In R. Behrens, D. McCormick & D. Mfinanga (Ed.), *Paratransit for African Cities.* Routledge.

Kyem, P. A. K. (2004). Of intractable conflicts and participatory GIS applications: The search for consensus amidst competing claims and institutional demands. *Annals of the Association of American Geographers, 94*(1), 37–57.

McHugh, B. (2013). Pioneering open data standards: The GTFS story. *Edited by Brett Goldstein with Lauren Dyson*, 125.

Ratti, C., Williams, S., Frenchman, D., & Pulselli, R. M. (2006). Mobile landscapes: using location data from cell phones for urban analysis. *Environment and Planning B: Planning and Design, 33*(5), 727.

Roth (2010). How google and Portland's TriMet set the standard for open transit data. Streetsblog. http://sf.streetsblog.org/2010/01/05/how-google-and-portlands-trimet-set-the-standard-for-open-transit-data/. Accessed January 5, 2014.

Sieber, R. (2006). Public participation geographic information systems: A literature review and framework. *Annals of the Association of American Geographers, 96*(3), 491–507.

Talbot, D. (2013). African bus routes redrawn using cell-phone data. *MIT Technology Review*. http://www.technologyreview.com/news/514211/african-bus-routes-redrawn-using-cell-phonedata/. Accessed October 6, 2014.

Townsend, A. (2013). *Smart cities: Big data, civic hackers, and the quest for a New Utopia*. New York: W.W Norton.

Townsend, A. (2014). Re-programming mobility: The digital transformation of transportation in the United States. http://reprogrammingmobility.org/wp-content/uploads/2014/09/Re-Programming-Mobility-Report.pdf . Accessed March 4, 2015.

Wakefield, J. (2013). Mobile phone data redraws bus routes in Africa. *BBC*, http://www.bbc.com/news/technology-22357748. Accessed September 30, 2014.

Wang, H., Calabrese, F., Di Lorenzo, G., Ratti, C. (2010). Transportation mode inference from anonymized and aggregated mobile phone call detail records.In *Intelligent Transportation Systems Conference* (ITSC 2010), pp. 318–323.

Williams, S., Marcello, E., & Klopp, J. M. (2014). Toward open source Kenya: Creating and sharing a GIS database of Nairobi. *Annals of the Association of American Geographers, 104* (1), 114–130.

Williams, S., White, A., Waiganjo, P., Orwa, D., Klopp, J. (Forthcoming). The digital Matatu project: Using cell phones to create an open source data for Nairobi's semi-formal bus system. *Journal of Transportation Geography* (Working Paper Submitted October 2014).

Wong, J. (2013). Leveraging the general transit feed specification for efficient transit analysis. *Transportation Research Record: Journal of the Transportation Research Board, 2338*(1), 11–19.

Woyciechowicz, A., & Shliselberg, R. (2005). Wayfinding in public transportation. *Transportation Research Record: Journal of the Transportation Research Board, 1903*(1), 35–42.

Zegras, P. C., Eros, E., Butts, K., Resor, E., Kennedy, S., Ching, A., Mamun, M. (2014). Tracing a path to knowledge? Indicative user impacts of introducing a public transport map in Dhaka, Bangladesh. *Cambridge Journal of Regions, Economy and Society*, rsu028.