Open Data for Inclusive Urban Public Transport Globally

Jacqueline M. Klopp, Felix Delattre, Antoine Chevre

For the French Development Agency
June 2019

Executive Summary

With the pressing need to address climate change and urbanization together, it is imperative to build inclusive, safe, resilient, and sustainable urban public transport globally. Yet significant barriers exist to achieving this goal including missing transport data and data infrastructure for public transport. The digital revolution and open data movement have important, yet underrecognized, significance for overcoming these barriers and improving public transport. This review provides the context for understanding why existing urban public transport data gaps and inequality problems present a serious but less recognized problem for affecting needed change in the public transport sector. It explores efforts to close urban transport data gaps with a particular focus on diverse African and Latin American cities where significant innovation is emerging. Through these cases we can gain insight into how critical challenges in closing the data gap are being overcome and how new data is helping to provide accessibility analyses, visualizations, planning tools and information services while moving a step closer to a "mobility as a service" paradigm. We also discuss how a "digital commons" approach based on open, collaborative data creation and sharing platforms networks and communities of practice can help scale these efforts and close data gaps while creating a new vision and set of tools for achieving mobility and access goals. This approach builds on the proven power of open standardized public transport data, open-source tools as well as knowledge sharing and building especially within and across cities enabling critical "data sovereignty" and local tech eco-systems that can improve transit planning and services. From this analysis, we provide recommendations on how to more effectively build "digital transport data commons" in the service of inclusive urban public transport globally and our sustainable development goals.

Table of Contents

Figures

1. Overview

1.1. The Urgent Global Agenda

Our planet is urbanizing rapidly while facing a critical climate change threat. We need to urgently move on to pathways that limit global warming to 1.5°C and to do so effectively "requires rapid and far-reaching transitions in energy, land, urban and infrastructure (including *transport* and buildings), and industrial systems." (IPCC 2018). Meeting the basic needs of swelling urban populations while addressing climate change, ensuring the integrity of critical ecosystems, and promoting economic productivity, public health and social inclusion is one of the major challenges of our times.

By 2050 two thirds of the world's population will be in cities. Cities will be where the battle against climate change and for sustainable development will be won or lost. This is why we now have an urban sustainable development goal to make "cities and human settlements inclusive, safe, resilient and sustainable" (USDG11). One of the key levers for addressing the inter-connected problems in cities involves improving urban mobility and access with mass public transport as the backbone of multimodal systems. To make this lever work we need to dramatically improve public transport. This will help avoid emissions intensive private car use and encourage retention of passengers in public transport. This requires increasing the level of service as well as expanding access and building inclusive multimodal transport systems.

Effective, well designed, inclusive urban public transport can produce "win-win" scenarios in terms of equity, public health, greenhouse gas reductions, economic productivity and access. This is especially the case in rapidly urbanizing and growing areas of the world like Africa and Asia. In many cities in these regions the majority of citizens already rely on public transport, often in the form of popular systems like minibuses, motorbikes and tuktuks along with non-motorized transport, mainly walking, but also bicycles.

Reflecting the importance of improving public transport and mobility to multiple development goals, we have a global target Sustainable Development Goal (SDG) 11.2:

"By 2030, provide access to safe, affordable, accessible and sustainable transport systems for all, improving road safety, *notably by expanding public transport, with special attention to the needs of those in vulnerable situations, women, children, persons with disabilities and older persons*."

In addition, it is widely recognized that transport is a key component in reaching other key SDGs around energy, gender equality, poverty reduction, decent work, climate and public health especially reduction of fatalities and harm from crashes and air pollution. As it can be seen in Figure 1 below from the Partnership on Sustainable Low Carbon Transport (SLoCaT 2016)[1].

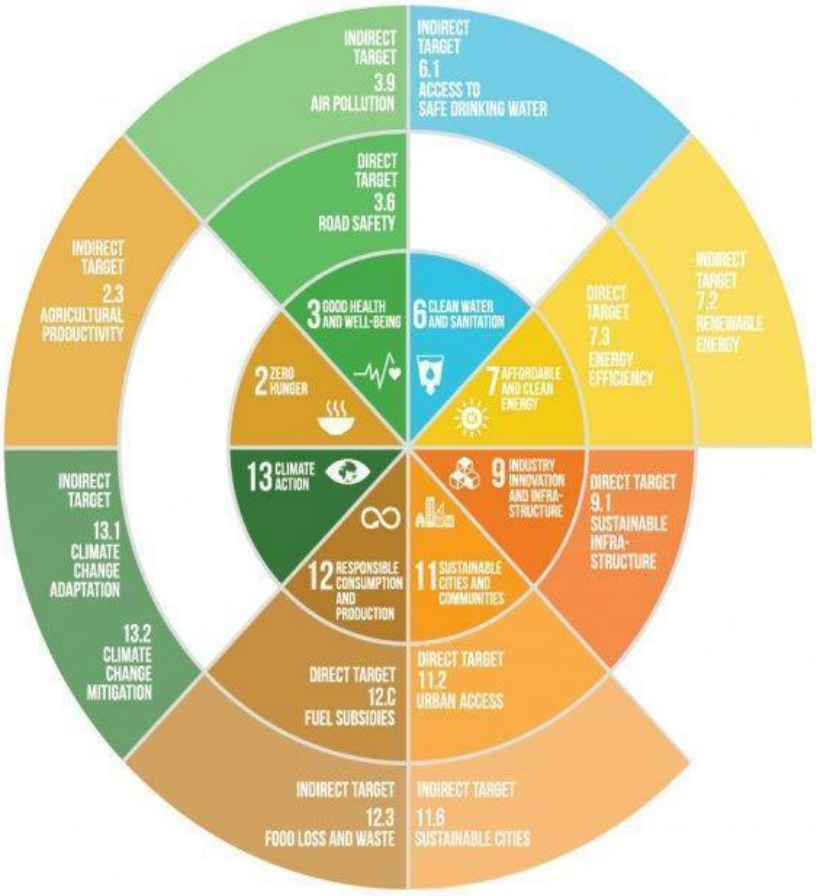


Figure 1: Transport-Relevant SDG Targets (SLoCaT 2016).

Currently, to address our key global climate targets and sustainable development goals, a number of critical global and regional partnerships have emerged. In 2009, the Partnership on

[1] Transport contributes *directly* to five targets on road safety (Target 3.6); energy efficiency (Target 7.3); sustainable infrastructure (Target 9.1), urban access (Target 11.2), and fossil fuel subsidies (Target 12.c) and contributes *indirectly* to seven SDG targets on agricultural productivity (Target 2.3), air pollution (Target 3.9), access to safe drinking water (Target 6.1), sustainable cities (Target 11.6), reduction of food loss (Target 12.3), climate change adaptation (Target 13.1), and climate change mitigation (Target 13.2)

Sustainable, Low Carbon Transport (SLoCaT) formed to "promote the integration of sustainable transport in global policies on sustainable development and climate change and leverage action in support of the implementation of the global policies". More recently, in 2017, the World Bank led the launch of "Sustainable Urban Mobility for All", a multi-stakeholder initiative focused on "universal, efficient safe and green" mobility (Figure 2).

Figure 2: The Sustainable Mobility for All goals

In order to track and assess progress on this critical agenda, the indicator for SDG 11.2, defined by the Inter-agency and Expert Group on the Sustainable Development Goal Indicators (IAEG-SDGs), needs to be monitored for progress. As it currently stands, this indicator is "the proportion of the population that has convenient access to public transport, by sex, age and persons with disabilities"[2]. As part of this effort, the SUMA initiative aims to track actual performance towards sustainable mobility in support of the 2030 Agenda and has set up an Elementary Global Tracking

[2] According to the latest UN Statistical Commission report access to public transport is considered convenient when an officially recognized stop is accessible within a distance of 0.5 km from a reference point such as a home, school, work place, market, etc. Additional criteria for defining public transport that is convenient include:
a. Public transport accessible to all special-needs customers, including those who are physically, visually, and/or hearing-impaired, as well as those with temporary disabilities, the elderly, children and other people in vulnerable situations.
b. Public transport with frequent service during peak travel times
c. Stops present a safe and comfortable station environment

Framework for Transport (GTF) and Global Mobility Report which is refined and updated every two years (SUMA 2017).

1.2. Missing Data, Transport Data Inequality and the Digital Revolution

With the growing call for action on well-defined targets around urban transport, it has also become increasingly clear that serious urban public transport data gaps and inequalities exist. This threatens our ability to analyze the current situation and trends and use indicators to monitor SDG 11.2. For many parts of the world, especially many cities in Africa, Latin America and Asia, accessible urban transport data including the most basic static information on routes and stops and service quality are scarce or missing. Often only informal knowledge of the public transport organisation exists, but this is not effective for planning for millions of people in metropolitan areas. One World Bank estimate suggests that 35% of the world's largest cities and 92% of the largest low and middle-income cities do not even have complete transit maps for planning or passenger information (Krambeck 2015). This has impacts on citizen passenger experience and overall effectiveness of public transport service provision including coordination between transport networks. Overall, without data, local authorities are planning without seeing their transport system as a whole.

An opportunity exists with the digital revolution to address this problem. Never before have we had such a plethora of tools to build initiatives and create or harness growing quantities of data to fill these gaps. The spread of new forms of data and technology, including the rapid expansion of cellphone use, offers opportunities to create both new services and also missing public transport data. In addition, better technology and economies of scale mean that it costs less than before to capture, store and process data (Goldsmith 2013).

Private sector actors (such as Google, Uber or Lyft) are increasingly collecting large quantities of data and are offering various transport and transport information services. However, in many of these private sector efforts, data and the tools and knowledge to use this data effectively within the transport sector are valuable commodities that are sold, often at high prices. In addition, the collection of data from users including through location-based apps is leading to serious privacy concerns (Valentino-DeVries et al. 2018). In places with weak or absent data privacy laws this is a growing and very serious area that requires attention (Nyabola 2018).

Besides this collection of user data, many companies like Google also rely on high quality non-personal base data about transit systems that they typically get from transit agencies themselves (Transit Center 2018). Whereas, in Europe, United States and East Asia, public authorities of large cities have more resources and skills to generate, manage and publish basic data which private companies use to produce useful routing and other services through apps, many African, Latin

American and Asian cities are increasingly dependent on data created, understood and captured by private sector actors. This limits what these cities can achieve. Indeed, some companies are offering to create data for these cities that are usually collected by transit authorities in other places. This creates asymmetry in power and understanding around data and the technologies used to generate and use this data.

This also raises potential data access and sovereignty issues. Recently, a blogger for the International Transport Workers' Federation, a network of 670 trade unions, in 140 countries, representing millions of working men and women in all transport sectors noted,

> *"…a city run on data has to collect that data. Who collects it, analyses it and owns it then becomes an incredibly powerful player in that city. This has potential impacts on democracy, on social and gender exclusion, and on all sorts of issues that will affect transport workers as the people that keep goods and people circulating across it, as well as inhabitants of these cities. Public transport workers, formal and informal, and citizens must be part of a democratic planning process on the future of cities, to ensure that data is a public resource."* (ITF 2018)

In this context, the importance of the open data defined as "data that can be freely used, re-used and redistributed by anyone" in the transport sector and the link to greater transparency in planning stands out clearly. This can be the motivation for nurturing inclusive data ecosystems or Communities of Practice around one city, or region to help produce, clean, update data and innovate around this data to provide new services. Supporting the functioning of a strong open data value chain of collection, publication, uptake, and impact is especially critical when a city has weak internal capacities and ecosystem dynamics.

Building communities of Practice (CoPs) defined as "groups of people who share a concern or a passion for something they do and learn how to do it better as they interact regularly" (Lave et. al. 1991, Wenger 1998, Wenger et. al. 2002) becomes particularly important in building data ecosystems and capacities in cities. A CoP requires (1) a shared domain, (2) a community pursuing their interest in their domain and strategically exchanging with others and (3) a practice that develops "a shared repertoire of resources: experiences, stories, tools, ways of addressing recurring problems—in short a shared practice" (Wenger-Trayner et. al. 2015).

Digital technology can play a key role in building such communities of practice; either as a space of gathering as virtual communities of practice (Dubé 2005) or as a leveraging element of the shared practice: the resources can be digital information and assets, the goals can be digital products, and the processes are often supported and defined by digital tools (like forums, social

media groups, email lists, project management software, Version Control Systems, etc.). A very prominent example of emerging CoPs are Open Data and Free and Open Source communities which organize over the internet, and their practices are usually driven by the programs and data sets they develop. Those communities are inclusive and open for people that share the goals and domain, want to be part of the community and agree to build on top of common goods and largely transparent practices.

Open Data and Free and Open Source communities by default are communities of practice and directly support the approach described in SDG 17 to "strengthen the means of implementation and revitalize the global partnership for sustainable development". In the context of global mobility, the specific targets on the topic of technology described in the SDG 17.6 and SDG 17.7 are particularly relevant:

> *Enhance North-South, South-South and triangular regional and international cooperation on and access to science, technology and innovation and enhance knowledge sharing on mutually agreed terms [...]*
>
> *Promote the development, transfer, dissemination and diffusion of environmentally sound technologies to developing countries on favourable terms [...]*

Communities of practice can help support SDG 17.6 and 17.7 and are beautifully complex and open-ended. They include multiple disciplines and all kinds of people, organizations, both non-for-profits and governmental entities, and can be effective ways to build data value chains and local knowledge around data in specific places. Thus, building communities of practice around open data is a strategic approach to addressing the data constraints to improving public transport globally.

1.3. The implications of public transport data gaps and inequality

Existing data gaps cause a number of serious problems for how urban public transport is planned and implemented in many of the world's cities. Specifically, without basic urban transport data, better design and policy along with needed evaluation and monitoring are difficult to do, leading to accountability and planning problems. Challenges exist in building strong structured frameworks for creating transport data including of base public transport data, roads characteristics and public transport supply and operations. Without this data cities miss the opportunity to improve planning as well as create value or invent services through communities of practice- as expressed in the SDG 17- around digital technologies and data for mobility.

Growing evidence, primarily from North America and Europe, is showing that open transport data and the eco-systems and communities of practice that emerge out of this open data are producing important innovation, benefits and transport improvements (APTA 2015, Schweiger et. al 2015, Coplaert and Rojas Melendez 2019). This means that countries and cities without public transport data- and specifically open non-sensitive base data- lose out on the new opportunities produced by technology for improving public transport, as well as monitoring these improvements.

- **Data gaps handicap public authorities in understanding, planning and regulating the transport/mobility sector**

Where there is no or little publicly provided transport services, local governments have no basic information on the current situation of public transport provision in their cities. One reason that authorities often do not have good data is that public transport is often privately provided and public transport authorities do not even exist. Information is then divided among a multitude of actors and is very heterogeneous, of unclear quality and often inaccessible. Public authorities are then "blind" to their system as a whole which makes it difficult to develop constructive dialogue with transport sector stakeholders to enhance the public transport and overall mobility system. With increasing and much needed investment in mass transit systems, data gaps mean implementing long and costly data collection from scratch usually surveys. Within this data collection, most often the dominant informal or popular transport is excluded from consideration and this is lost opportunity to engage this critical sector through mapping (Klopp and Cavoli 2017, 2019). Instead, mass transit is often considered as a means to replace informal transport even when it is increasingly clear that such systems should be integrated for high connectivity and access (Klopp and Cavoli 2019, Stokes 2019). Data is needed to understand and *see* popular transport – their networks, dynamics and importance. Without this ability to visualize and analyze these systems, decision making processes for transport are, in effect, involving the body of city without understanding the vast, dynamic circulatory system that gives the city life.

- **Critical passenger information systems are not possible without data**

Providing high quality passenger information is a key form of access. Well- functioning multi-modal public transport systems involve providing passengers with basic transit maps, transit signs, screens and apps –information systems that allow them to better plan efficient trips by public transport and avoid unnecessary waiting. Growing evidence suggests the importance of this information, especially open and real time data, for public transport users and systems.

This is especially true as we move closer to a "mobility as a service" (maas) paradigm. The idea of maas is that mobility becomes a service, and people do not need to own personal transport modes especially cars. Maas requires that data is available on where and when services are available, open ticketing and high levels of integrated multi-modality (Colpaert and Roja Melendez 20190. This enables users to access transportation services from public and private providers through a unified platform (an app) that creates and manages seamless multimodal trips, which users can pay for with a single account.

Whether for building mobility as a service or improving public transport immediately, real time passenger information is critical for seamless travel across many modes. Overall, satisfaction and use of public transport can be increased by improving information systems, because this information empowers users to make better decisions, wait less and make more efficient trips (Shaheen et al. 2016). Growing research suggests that providing real time data can increase ridership (Tang and Thakuriah 2012, Brakewood et al. 2015, Brakewood et al. 2011), make transit feel safer (Ferris et al. 2010, Transport Research Board 2013, Klopp et al. 2015). In the longer term, this data is a requirement to implement Mobility As A Service (MAAS) where people can choose the best mobility option. Integrated fare systems are also critical for passenger ease and moving towards (MAAS) and can provide useful data for planning.

In sum, high quality information across public transport networks and modes is a means to offer more mode and trip choice. Ultimately, the high level of cost-effectiveness, convenience and flexibility of travelling on a data enabled seamless integrated multi-modal system with mass transit as the backbone effectively challenges the idea of the car as the most liberating way to move within a dense urban environment. This is a key and potentially powerful lever for reaching sustainable development target 11.2 and air pollution and GHG reduction goals as well.

- **Without data communities of practice, a local transport technology sector cannot emerge and opportunities for value creation and employment are lost**

Open data creates an opportunity for technology companies, researchers, and all kinds of entrepreneurs to develop innovative products of all kinds and provide transit users, operators and authorities with new services. This process generates good jobs for young professionals and transfers knowledge across cities. It is essential to see the value that can be created on top of open data. Publishing data especially in standardized formats generates significant impact (Open Data Watch 2018).

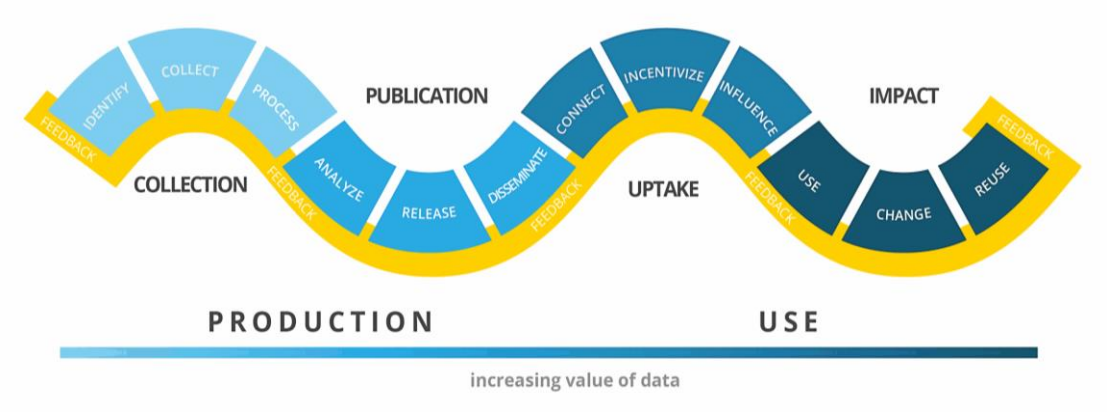


Figure 3: The data value chain (Source: Open Data Watch 2018)

Open transport data can thus help contribute to increased economic development as well as resilience in transport planning and service provision. It allows entrepreneurs in the region to build up new businesses that improve services and planning as well as researchers and analysts to develop new, locally specific knowledge to guide decision-making. Recognizing the importance of these benefits, the international development community has started to target Open Data as one of their goals for sustainable development (Principle "Use Open Standards, Open Data, Open Source, and Open Innovation" and principle "Reuse and Improve" of the widely accepted "Principles for Digital Development") (Principles for Digital Development 2015).

- **Data gaps will put in jeopardy our ability to monitor and promote the 2030 Sustainable Development Goal Agenda**

Currently, we do not have data and stream-lined methods in place on the commonly agreed indicator for SDG11.2 (See Figure 4).

Figure 4: Current status of SDG Indicator 11.2.1 (UN Statistical Commission Accessed May 2019)

The UN classifies this indicator as under Tier II, meaning the indicator "is conceptually clear and an established methodology exists but data is not easily available."[3] Further challenges include the fact that some aspects of the data will require collecting at the municipal/city level and thus data creation and maintenance capabilities at this level need support. Quite simply, base data on many public transport systems including basic public transport points are needed but missing.

1.4. The proven power of the digital commons approach

Given the value that building open transport data can add to global policy efforts to improve inclusive public transport, it is critical to find ways to scale up the creation, sharing and use of critical data. Currently, transport data is being produced and used by a wide array of actors but not in any standardized format, with little sharing. Hence, the potential value of this data is widely lost.

Especially in cities and regions without strong transit agencies and policies, many different data producers and consumers often work in parallel, and data is poorly shared and leveraged for improving public transport. Given the importance of data as a common resource, critical infrastructure and an asset for shared planning and action for cities, such data should be widely shared and open. Ideally, existing efforts would join together, enable ecosystems and foster vibrant communities of practice. This "digital urban transport data commons" would increase efficiencies across a network of actors aiming to improve public transport and integrated, holistic planning (as described in SGD 17), as well as help spark shared dialogue, accountability and monitoring including in support of SDG11.2 and 17.6 and 17.7. Overall, evidence is increasing that open public transport data can have important positive impacts and is the building block for moving into a new mobility as a service paradigm.

1.5. Leveraging Open Public Transport Data

The case for *good, non-sensitive open base data on public transport* especially routes, stops, service status, fares, times *as well as* planned investments and network changes along with real time information is growing. More and more transport authorities and companies, especially in

[3] The actual and recommended data sources for this indicator are the following: - Data on the location of public transport stops in a city: city administration or service providers, GIS data - Dwelling units within 500m of public transport stops: Census, GIS data - Number of residents per dwelling unit: Census/household survey - Household surveys that collect information on the proportion of households that declare they have access to public means of transport within 0.5 km. These surveys can also collect information about the quality of the service.

North America and Europe, but also increasingly in the Asia-Pacific region, are making base public transport data open, especially in the standardized General Transit Feed Specification (GTFS) format (Rojas 2012, APTA 2015, Schweiger et. al 2015, Hogge 2016, Harmony and Gayah 2017, Transit Center 2017). Where transit authorities are weak or do not exist, open data communities have also started to use the new technology to create GTFS data sets and make them available.

Open data is defined as "data that can be freely used, re-used and redistributed by anyone - subject only, at most, to the requirement to attribute" (ODI 2018). The Open Data Handbook summarizes the key attributes of open data as:

- **Availability and Access**: the data must be available as a whole and at no more than a reasonable reproduction cost, preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.
- **Re-use and Redistribution**: the data must be provided under terms that permit re-use and redistribution including the intermixing with other datasets.
- **Universal Participation:** everyone must be able to use, re-use and redistribute - there should be no discrimination against fields of endeavor or against persons or groups. For example, 'non-commercial' restrictions that would prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

(ODI 2018)

"Good" open data according to the Open Data Institute should be 1) linked to the internet so that it can be easily shared and talked about 2) available in a standard, structured format, so that it can be easily processed 3) have guaranteed availability and consistency over time, so that others can rely on it, and 4) is traceable, through any processing, right back to where it originates, so others can work out whether to trust it (ODI 2018).

The United States National Academies of Sciences, Engineering and Medicine surveyed sixty-seven transport agencies across the globe on their experience with open data and found evidence that significant benefits emerge from making public transport data open. Specifically, the report notes:

- *The impacts of open transit data on customers and the general public are significant;*
- *The impacts on the private sector have been encouraging over the past several years. Applications and visualizations that could not necessarily have been conceived or developed by a transit agency have been created;*
- *The legal fears often thought to be barriers to opening transit data have not been realized;*

- *Standards greatly facilitate the use of open transit data, although this sometimes requires additional effort in producing the data;*
- *Engaging with data users and re-users has the potential to increase the value of the applications and visualizations* (Schweiger et al. 2015)

Seventy eight percent of agencies also reported that the public was more aware of public transport services because of open data (Schweiger et al. 2015).

Increasingly, encouraged by these positive impacts, transport agencies are creating "developer portals" with "information about the different types of data provided, licensing information, and contact information where users can ask questions along with links to interesting projects created using the data" (APTA 2015). Opening public transport data allows many actors to leverage the data better for improving planning, service quality and passenger information and even allows for public input into correcting or cleaning and hence, improving the data itself. It also creates more transparency and understanding (Cerrillo-I-Martinez 2012, Kitchen 2013)

Open data also allows third party innovation especially around passenger information systems. This is most evident in the "an ecosystem of third-party apps being developed for the dominant smartphone platforms" (O'Brien 2017). Open public transport data thus can also help foster business development, and some app developers are able to generate income from innovative apps they create using open data. Indeed, a key factor in Transport for London's (TfL) decision to move towards an open data strategy was the fact "that developing apps in-house that served every smartphone platform—would have been an expensive undertaking" (Hogge 2016). Allowing the private sector to develop diverse quality products on top of the data can be highly cost effective. TfL's Head of Bus Systems & Technology calculated that apps powered by TfL's bus data will deliver £83m of customer benefit over 10 years, at a cost to TfL of £820,000 (Reed 2015 Cited in Hogge 2016). TfL had over 5000 registered developers, 362 apps available powered by TfL data reaching 4 million people with an estimated £15m-£58m value of time saved by users of these apps in 2012 (Hogge 2016).

Some of these companies are also able to generate more data as apps gain users who, in turn, feed back into data creation and, in some cases, even transit provision. For example, CityMapper, a London-based company, used its data to discover demand for missing bus links and launched a bus-taxi hybrid services on these routes (Hern 2018). Transit Screen, Inc. uses open data from public transit agencies to develop informational screens for building lobbies and public spaces. Overall, many of these businesses develop services that make shared mobility and public transport more attractive. These efforts help build a rich eco-system in support of the global agenda of improving public transport and reducing individual car use. The evidence speaks to the

power of developing ecosystems and communities of practice around open transit data. This approach also shows how to reach SDG 17.6 and 17.7 targets by making technology available at fair terms and encouraging knowledge transfer and learning across cities.

Despite growing evidence and supporting theory around these benefits, the process of moving from a closed to an open data approach is not always straightforward and requires a change of bureaucratic culture as well as overcoming a number of barriers (Rojas 2012, Colpaert et al. 2017). Challenges include risk-averse institutional cultures inside transit agencies and city government more generally, privacy and legal concerns, proprietary vendor contracts that preclude sharing data with third parties, and time-consuming technical efforts to produce accurate datasets suitable for public disclosure and use. Increasingly, government's problem is also not about capturing the data, but rather "having sufficient resources to clean and analyze the information in order to address issues, improve performance and make informed decisions" (Goldsmith 2013).

Many agencies provide short and simple license agreements in order to encourage users. Data users usually do not have to sign an agreement or provide anything to the agency – downloading the data is confirmation of agreement with the license. According to the US National Academies report common elements of public transit agency data licensing agreements include

- The agency reserves the rights to its logo and all trademarks.
- The data are provided without warranties.
- No availability guarantees are expressed or implied.
- The agency retains full rights to the data.
- The license is free of charge and is made between the [agency] and a licensee.
- The licensee may freely copy and deliver; modify and use (e.g., for commercial purposes); and combine and use as part of an application or service.
- The name of the licensor (e.g., Finnish Transport Agency) should be shown.

(Schweiger et al 2015)

It should be noted that while open static data raises few legal problems, concerns do exist around privacy, especially around making real time data open (Scassa, T., & Diebel, A. 2016). More work needs to be done on developing data protection laws and policies as well as increasing public awareness around privacy protection issues especially in contexts where such laws are currently weak or underdeveloped (Nyabola 2018).

1.6. Leveraging Common Data Formats: General Transit Feed Specification

One critical feature of open transit data is the possibility of sharing and using this data for passenger information systems and urban analyses. Existing public transport data standards have been overwhelmingly complex in the past (e.g. Transmodel), or are relatively recent standards (like NeTEx, formally Network Exchange *PD CEN/TS 16614-1:2014', PD CEN/TS 16614-2:2014 and PD CEN/TS 16614-3:2014*). For this reason, in 2005 a simpler format oriented towards journey planning - the Google Transit Feed Specification (GTFS) was created out of a collaboration between technical managers of TriMet, the public transit agency in Portland, Oregon and Google developers (McHugh 2013). In 2009 it was renamed the General Transit Feed Specification. Despite not being an official standard, it has become the de-facto standard, the most used and most common format used by transport authorities in most places of the world. It is designed for data integration and interoperability as well as the use of a wide array of open source tools, saving time and money for agencies and allowing for more transit comparability across cities (Wong 2013).

The GTFS format involves a set of simple related csv (comma separated values) files with each file capturing a particular aspect of related transit information: stops, routes, trips, and other schedule data. This data structure was chosen because "it is easy to view and edit using spreadsheet programs and text editors", and it was important to make it easy for agencies to edit and hence to allow them to participate (McHugh 2013). A GTFS-Real Time extension was later developed in 2012 to address the need for real time information for users and not just static data.

Filename	Required	Defines
agency.txt	**Required**	One or more transit agencies that provide the data in this feed.
stops.txt	**Required**	Individual locations where vehicles pick up or drop off passengers.
routes.txt	**Required**	Transit routes. A route is a group of trips that are displayed to riders as a single service.
trips.txt	**Required**	Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.
stop_times.txt	**Required**	Times that a vehicle arrives at and departs from individual stops for each trip.
calendar.txt	**Conditionally required**	Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available. This file is required unless all dates of service are defined in calendar_dates.txt.
calendar_dates.txt	**Conditionally required**	Exceptions for the service IDs defined in the calendar.txt file. If calendar.txt is omitted, then calendar_dates.txt is required and must contain all dates of service.
fare_rules.txt	Optional	Rules for applying fare information for a transit organization's routes.
shapes.txt	Optional	Rules for drawing lines on a map to represent a transit organization's routes.
frequencies.txt	Optional	Headway (time between trips) for routes with variable frequency of service.
transfers.txt	Optional	Rules for making connections at transfer points between routes.
feed_info.txt	Optional	Additional information about the feed itself, including publisher, version, and expiration information.

Figure 5: Overview of the GTFS format's files and contents

The common format and the open specification created the base for actors including transport agencies, local start-ups, Open Source communities, research institutions and enterprises like Google to build momentum around digitalization in mobility. A successful community of practice is now transforming the way digital devices are supporting public transport especially through providing real-time data. Importantly, this community is also speeding up the creation, use and sharing of public transport data in many cities where this data never existed.

Current efforts are also underway to explore how to further improve the GTFS format in light of a number of challenges including 1) the presence of on demand transit 2) measuring actual accessibility for those who are unable to walk 3) incorporating unplanned service changes and 4) addressing fare integration (modelling fares can be complex especially where fares are unregulated). To address some of these issues, a new more flexible GTFS-flex is being developed and tested particularly by a team called Datamobility.org incubated at the Rocky Mountain Institute in the United States (Transit Center 2018). These modifications will require input from communities of practice around the data to ensure they work for diverse public transport and mobility systems.

2. Digital tools for Inclusive and Integrated Public Transport

Since its launch in 2005 with Portland as the first city providing GTFS data to Google Transit, this de facto GTFS standard is now widely used and the digitalization of mobility continues to spread across the world, especially the US, Europe and parts of Asia. However, barriers appear to exist to developing standardized transit data for the rapidly growing cities of Africa, Asia and Latin America. Without standardized data these cities will miss critical opportunities for better planning, infrastructure design, monitoring, local economic development and information systems.

In many cities, one reason data does not exist or is inaccessible is that public mobility is dominated by minibuses, sometimes called paratransit because they do not have fixed schedules and routes and stops can also vary. Many operators in paratransit are small businesses that may not see the utility of data or do not have the time and money to collect it. Sometimes, the fact that some of what these businesses do is informal means that they also may also wish to stay under the government or public radar. Another reason for lack of data is that governments, used to seeing these systems as "chaotic" or too complex to address, often do not bother to require operators-including the larger businesses-to collect and share data. Worse, some government and industry actors collude and mutually benefit from the lack of transparency.

In addition, transport authorities often do not exist or are weak in these cities and even when government agencies do collect data, they most often hire consultants and do not always make the data collection methodology and data collected by these consultants accessible and open. Little institutionalization of local data creation capacities tends to occur within technical assistance which is also consultant driven. In some cases, as in India, transport officials also wish to find ways to monetize data and hence resist opening it up (Abisla 2019). Civil society or private sector initiatives are either not taken into consideration by local transport officials or cannot be relied on, because critical data and infrastructure are left to the control of private companies who do not share data. Finally, the flexible and demand responsive nature of these systems make some features like routes variable which needs addressing within data collection and standards. This is an intrinsic challenge to paratransit data collection and the GTFS standard needs modification because it was originally designed for scheduled, more fixed systems.

Despite these challenges, new possibilities have emerged to creatively use GPS enabled mobile phones to collect data at a low cost in cities that are currently missing valuable transport data. Successful efforts have created GTFS data in African, Asian, Middle Eastern and Latin American cities, starting to fill critical gaps and inequalities in our global transport data. These efforts are

also bringing together diverse actors to build coordinating networks and learning communities for data-driven approaches to improving public transport.

2.1. Proof of Concept: Case Studies from Kenya, Nicaragua and Ghana[4]

To illustrate how high quality transport data can be created and used for improving transport in cities that have high levels of paratransit, we look at three "proof of concept" projects in Kenya, Nicaragua and Ghana. All three emerged within the context of international cooperation and, by making data openly, available built communities of practice around the data, city and the people and organizations involved in data creation and innovation efforts. All three projects created the data in GTFS format, designed a paper map and brought the data into passenger information systems in some form on the web and on mobile phones.

The **DigitalMatatus** project is one early mapping project that helped catalyze ongoing mapping work in African cities. This collaboration between the University of Nairobi, Columbia University, MIT and a small design firm Groupshot mapped out Nairobi's minibus (matatu) routes and stops using GPS enabled mobile phones (Williams et al. 2015, Klopp et al. 2015). A schematic paper map was designed, and the data has been integrated into Google's services, providing comfortable routing on their website and application.

The **MapaNica.net** Managua Bus Mapping project was a community based data collection project. It was the kick-off project for the Nicaraguan OpenStreetMap community, in which over two-hundred students, professionals, public workers and companies were involved, to learn and to collaboratively collect the data of the capital's public transportation system. For the first time, they gave names to all bus stops in the capital, designed and distributed a paper map, created passenger solutions (such as mobile apps and websites) based on Open Source Software and made them available for the public (Dobush 2016).

Accramobility, a collaboration between the French Development Agency and the Accra Metropolitan Authority is another successful effort to build GTFS data for Accra's minibus (trotro) system (Saddier et al. 2016, Saddier et al. 2017). This project started with an institutional focus and later extended to involve citizens and collect data on the whole minibus system with the local OpenStreetMap community. The data was converted to GTFS and a paper map has been designed, further the project used existing open source solutions to provide a mobile app and website for passengers in Accra.

[4] For more Indian Cases see Abisla 2019 and for more African cases see Klopp and Cavoli 2019. For Cairo see transportforcairo.com.

2.2. New Visualizations of Networks

Projects: DigitalMatatus, MapaNica.net, Accramobility.

All three projects introduced - for the first time - a comprehensive map of the public transportation system of their city: Accra, Managua and Nairobi.

The DigitalMatatus project with the Civic Data Design Lab taking the lead, designed a schematic map for 120 routes of the dominant matatu system prepared for print media. The map building involved local actors who reviewed key landmarks and matatu drivers who commented on routes and stops. The map is now in the process of being updated with matatu sector feedback.

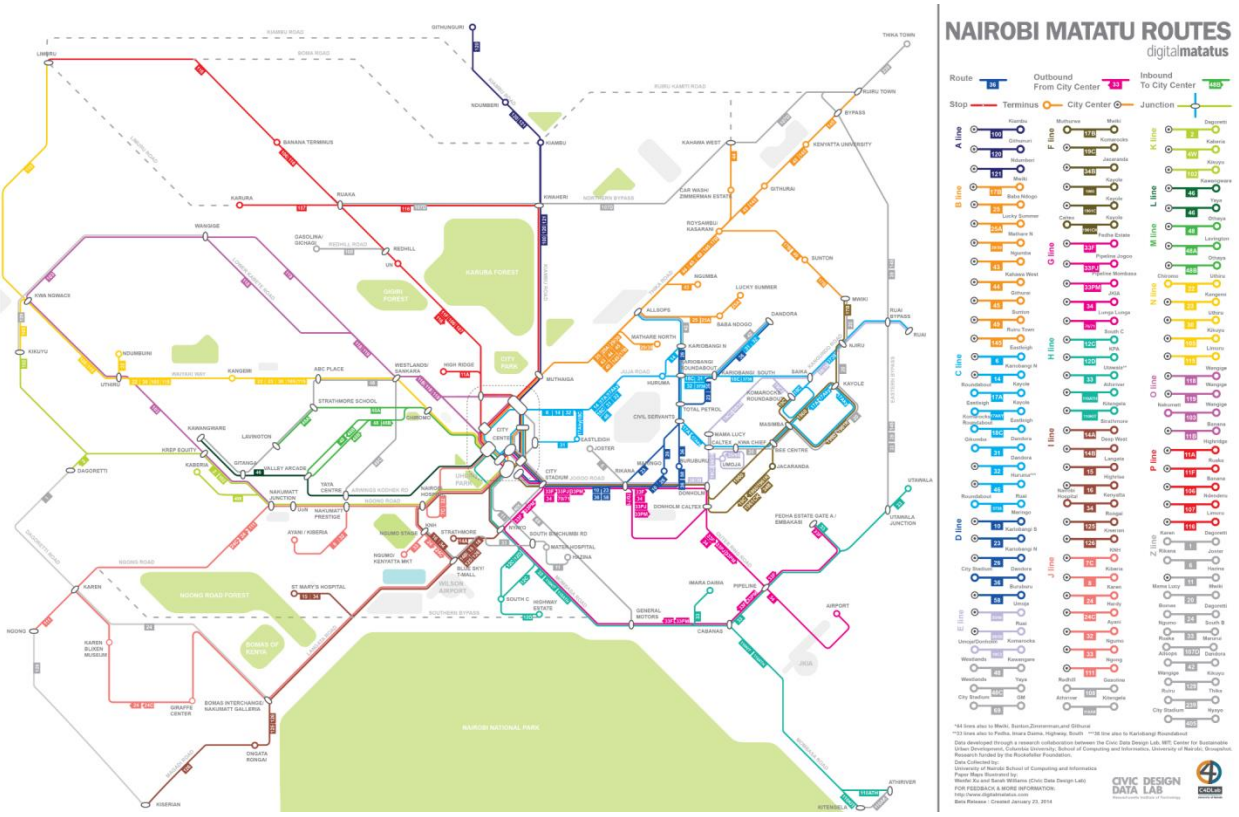


Figure 6: DigitalMatatus Matatu Map 2014

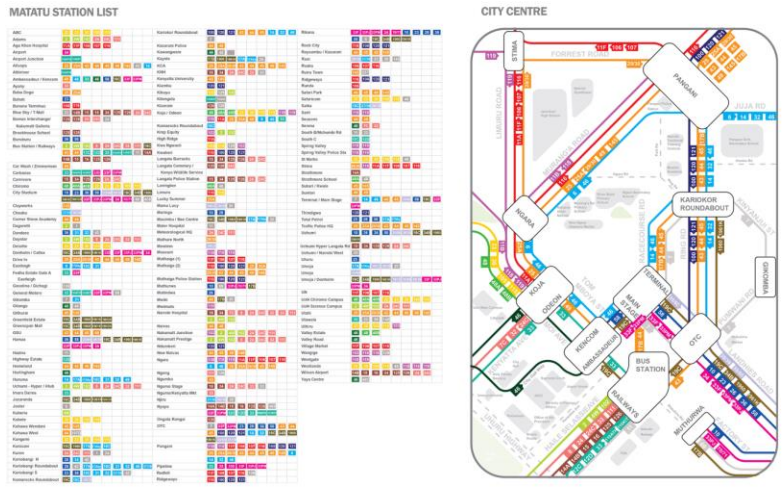


Figure 8: Overview of Nairobi's major transfer stations

The visualization started a wide process of reflection about the public transport system including an analysis of its network features. It is a strong starting point for emerging communities of practice.

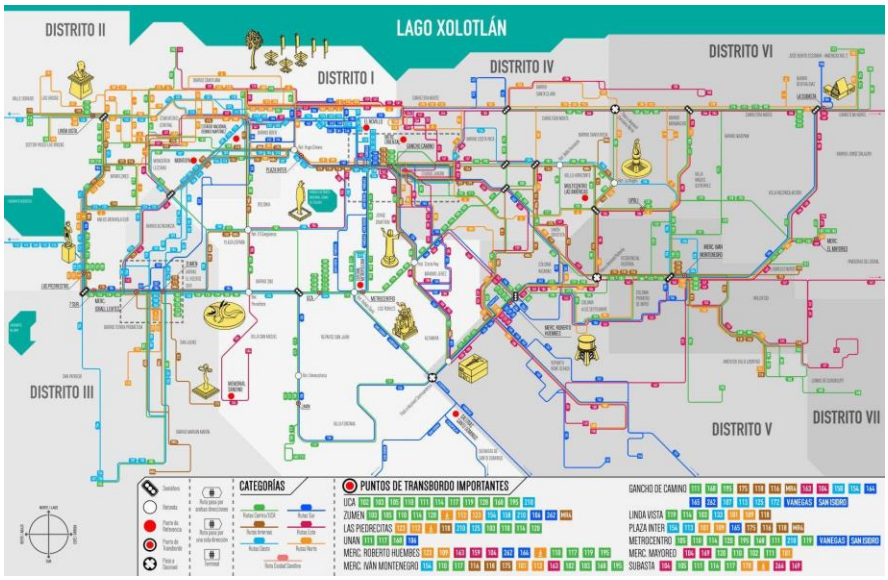


Figure 9: Managua public transport map 2016

The work of DigitalMatatus inspired other mapping collaborations and communities, like the Nicaraguan MapaNica.net to create a similar map reflecting the reality of their transport. In Nicaragua, the focus was put on enhancing the map with local style and symbols.

The visualization shows that there are fundamental differences between Nairobi, where a radial system prevails and Nicaragua which is more dispersed. In Nairobi all the key termini are in the center of the city. According to Jarrett Walker, this phenomenon "is a common thing that goes wrong in privately evolved systems. Every *matatu* wants to go downtown because it's the biggest market, and a matatu driver doesn't have to be coordinated with anyone else to fill a bus going to and from there" (2014). In contrast, in Nicaragua a network of buses was publicly coordinated, which led to a more distributed system. However, without network and passenger data, in both systems there has not been a focus on optimizing the network for higher efficiency and improved service based on available capacities.

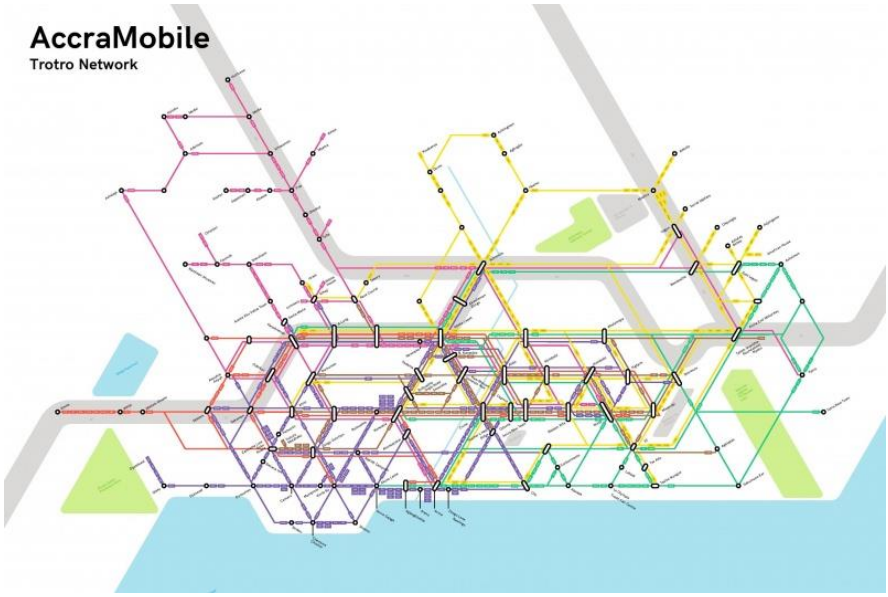


Figure 10: Schematic map of Accra's trotros 2017

Concordia University, which was a partner of the initiative Accra Mobility, developed a trotro map based on the GTFS data set. This map was drawn based on the different lines but requires work to improve readability and incorporate landmarks.

2.2. Passenger Information Systems

Projects: DigitalMatatus, MapaNica.net, Accramobility.

Beyond creating tools for understanding, seeing and refashioning the city, this new data creates a very important basic service for citizens; passenger information. Passengers can look at a map or on a transit app and see how to get from one place to another using the existing public transport system.

Passenger information systems improve wayfinding, helping people to plan more efficient trips and, when coupled with real time information, reduce waiting. This, in turn, improves the way passengers interact and feel about public transport and has a direct impact on accessibility to jobs and other opportunities. Actual improvements in service and infrastructure are still necessary, of course, but building these systems, especially using a participatory approach, opens the way for citizens to give feedback and helps to advocate for these improvements. The table below gives an overview of some of the passenger transit apps used with the data in our three examples (Figure 11). Note that with open data, multiple apps can be made available giving the user choice and with no cost to the city.

Name	Cellphone application	Web	Open Source	Comment	Projects
Bussi[5]	No	Yes	Yes	Very simple application. No routing.	MapaNica.net, Accramobility
Google Transit[6]	Yes (iOS & Android)	Yes	No	Market leader	Digital Matatus
OpenTrip Planner and Digitransit[7]	No. Only Responsive design	Yes	Yes	Widely used solution.	Not applied in any pilot project. But TriMet (Portland), and many transport agencies use OTP with their own interface

[5] https://gitlab.com/opentransitmap/bussi (retrieved 24.03.2019)

[6] https://transit.google.com (retrieved 24.03.2019)

[7] https://www.opentripplanner.org (retrieved 24.03.2019)

TransitApp[8]	Yes (iOS & Android)	No	No	Crowd-sourced real-time information, based on their users	MapaNica.net, Accramobility DigitalMatatus
Navitia and Transportr[9]	Android only	No	Yes	Great opportunities	MapaNica.net, Accramobility

Figure 11: Overview of passenger information applications

[8] https://transitapp.com (retrieved 24.03.2019)

[9] https://transportr.app (retrieved 24.03.2019)

Bussi - Bus Simple and Static Information System
Website: gitlab.com/opentransitmap/bussi

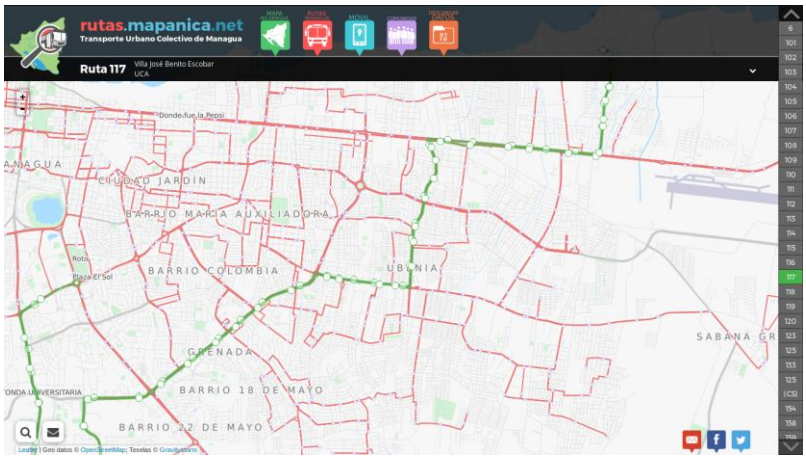


Figure 12. Bussi - Bus Simple and Static Information System

Bussi is a simplest website which is built to be run with lowest maintenance and running costs. It has been developed for the MapaNica.net Managua Bus mapping project and then picked up by Accramobility. Visitors can select routes and inspect their trajectory. It allows a location search (geocoding functionality) and importantly, it includes a functionality for passengers to provide immediate feedback about the correctness of the data via an email form. The Software has been Open Source from the beginning and is looking for an initiative to develop it from a particular solution into a product, to be used by everyone.

Google Transit
Website: transit.google.com

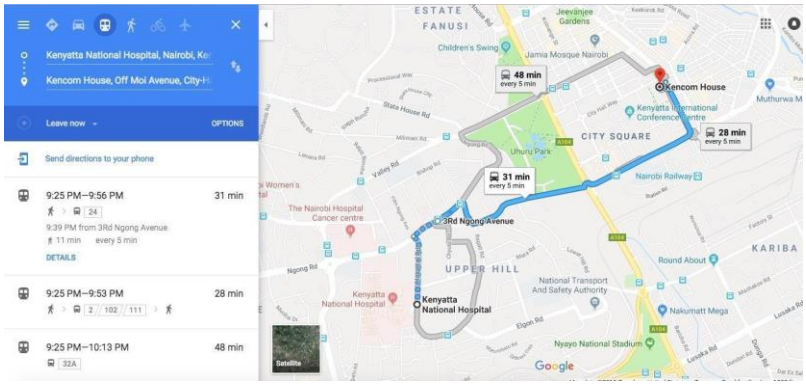
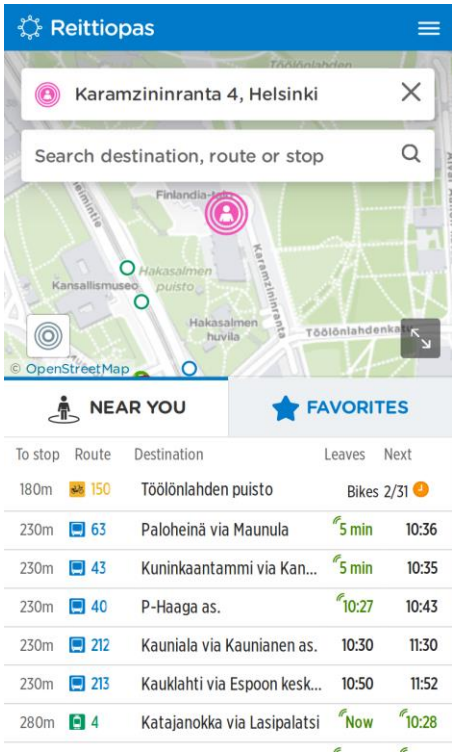


Figure 13: Google Maps view in Nairobi

Google Transit is the market leader and the apps are famous and well-integrated with cellphones and websites. DigitalMatatus' data is on Google Maps allowing trip planning. The data gets on average 300,000 queries a week suggesting demand. The application is full-service, no further participation nor responsibilities are involved although users can send feedback to Google. Often these services are attached to exclusive agreements, although this is not the case for DigitalMatatus data. The challenge is to keep the data updated given the highly dynamic nature of Nairobi.

OpenTripPlanner and Digitransit

Websites: opentripplanner.com, digitransit.fi

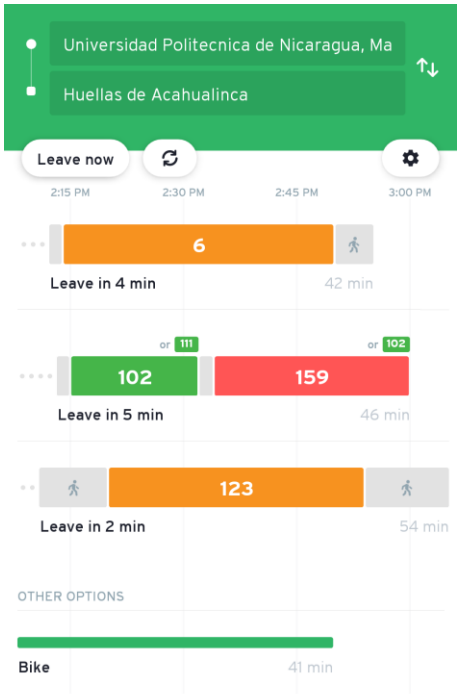


The OpenTripPlanner.com started in Portland. While it has not been used by any of the projects, it has a great potential because of its experience and maturity. It provides an easy-to-setup and though rich routing of GTFS. Together with the User Interface of Digitransit, an Open Source initiative which grew out of the Helsinki transport agency, it is a very promising solution, which allows comfortable access over web browser on computers and tables, but it is also optimized for mobile screen resolutions. It includes functionality for real-time data and it is being used by transport agencies all over the world. The OpenTripPlanner is the most versatile Open Source passenger information system available. There is a living ecosystem around it.

Figure 14: Digitransit, the Open Source Interface for the OpenTripPlanner used in Finland

TransitApp

Website: transitapp.com

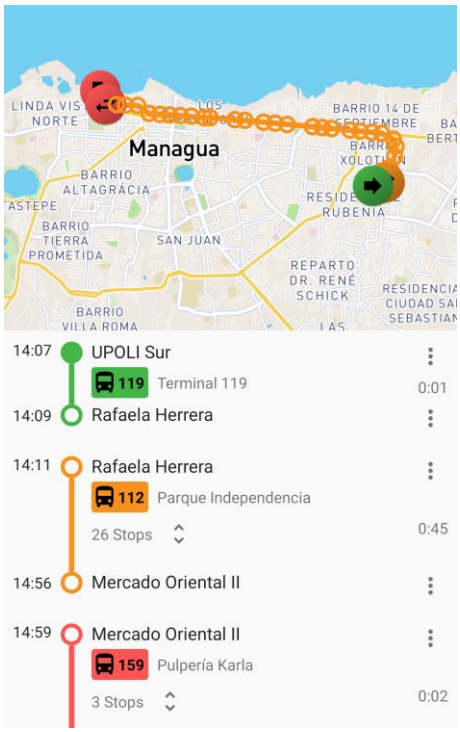


TransitApp is a commercial application that focuses on the "Go" mode based on the perspective on the passenger that wants to take a bus nearby. It collects data on their users to match it with schedules and provide real-time data in places where there is no formal system in place. The company has been very open to collaboration with open datasets from developing countries. The data of Managua and Accra has been integrated into their Android and iOS application and with its crowd-sourced real-time data approach, TransitApp provides the most accurate information on public transport systems in those two and other cities.

Figure 15: TransitApp - Go mode

Navitia and Transportr

Websites: navitia.io, transportr.app



Navitia is a strong routing engine for public transport, developed by the French company Kisio. It can be used by any kind of application and most consumers of this service implemented their own and closed solution.

There is one solid Android application - Transportr - which is providing transport routing for many cities all over the world. It is Free and Open Source Software and announces that it respects the users' privacy. Transportr is the application of choice of the MapaNica.net and Accramobility projects. It supports real-time data even though this data is not yet available for Managua or Accra. Because of its Open Source nature and its openness to user contributions it is adaptable. It is a solid starting point for cities, companies and entrepreneurs that want to build an application for their public transport network.

Figure 16: Transportr Android application

It is important to note that none of these applications are a substitute for proper information systems including on street information screen, signs and service that take into consideration language and culturally specific ways that people navigate. Much more experimentation is required to build well-functioning formal information systems for passengers in most of the world's cities where they do not yet exist.

2.3. Accessibility Studies

Projects: DigitalMatatus.

When transport data is overlaid with other kinds of data, we can explore how a mobility system generates *access* to services and opportunities in a city. It is really *access* we wish to improve not just mobility. With the rise of GTFS data sets, we can now overlay this data with land-use and jobs data to measure accessibility and see how different configurations of transport networks improve or diminish access to services and opportunities of different groups of people. This is critically important for bringing equity into planning. The World Bank, for example, was able to

use DigitalMatatus data to explore *physical* access to hospitals in Nairobi generated by the matatu system.

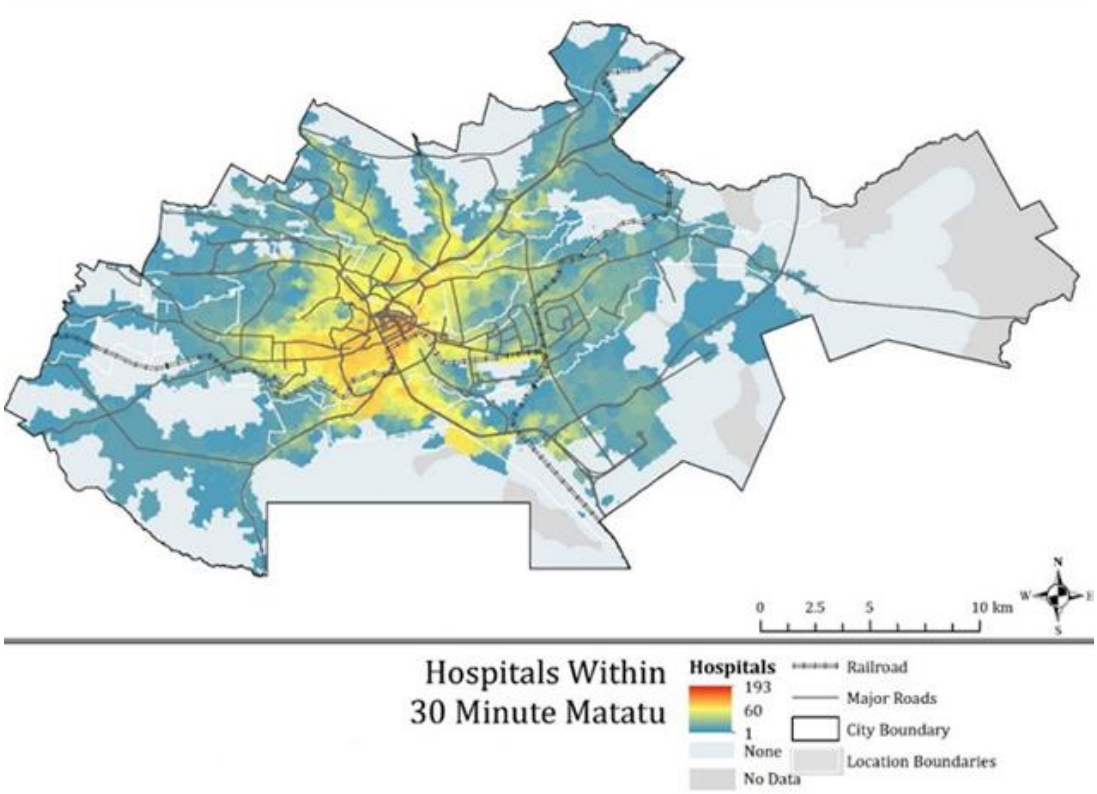
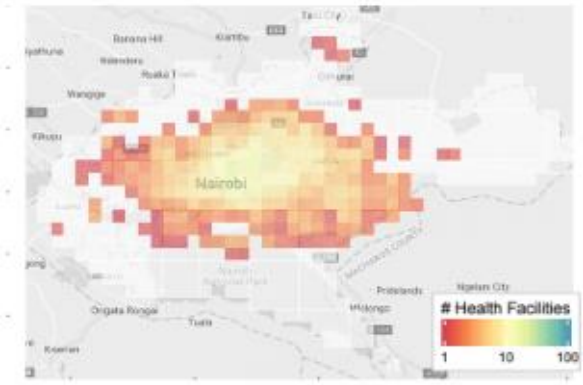


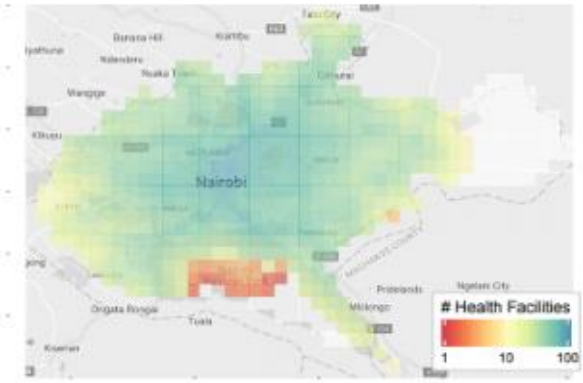
Figure 17: Hospitals Within 30 minute bus ride (World Bank 2016)

Good physical access is clustered in the center of the city. Disturbingly, spaces exist in the city where people do not have a hospital within a 30 minute or 60 minute matatu ride. This visualization also shows the ways that transport and land-use are inter-related; the problem of physical access to hospitals in Nairobi can be solved by building more dispersed (and ideally affordable and well equipped and staffed) facilities as well as encouraging more and better housing in well-serviced areas- by redesigning matatu routes or by some combination of these interventions depending on a number of important other factors.

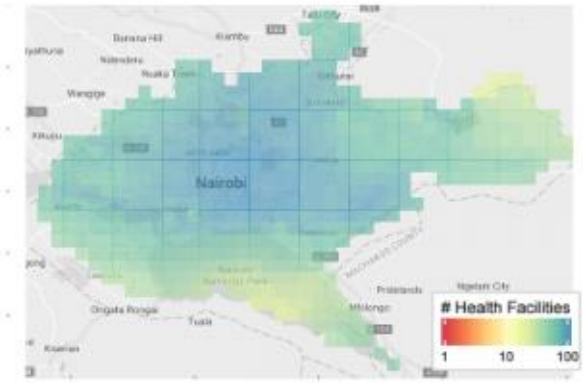
While seventy percent of Nairobi's adult residents use matatus every day, many of the city's poorest residents can only afford to walk. In contrast, the wealthiest residents travel by personal vehicle. Thus, by comparing access by mode, we can get a crude idea of the inequality of physical access to the city by socio-economic status. Below you see a visualization of access to points in the city by walking, minibus (sometimes called paratransit) and car.



(c) Walking Gravity ($G_i^{walking}$)



(f) Paratransit Gravity ($G_i^{transit}$)



(i) Driving Gravity ($G_i^{driving}$)

By comparing the swathe of blue colour you can see how cars give widest access to the city, matatus create substantial access and walking provides limited access to health facilities in the city. Again, this access is greatest in the center. This suggests why people may accept lower housing standards and high rents in centrally located slums; it allows more critical access to city services by walking and shorter and therefore cheaper matatu trips (Campbell et al. 2019). Similarly, using the DigitalMatatus data along with land-use data another World Bank study was able to show Nairobi's spatial mismatch in Nairobi (Avner and Lall 2016).

Overall, this means high quality and affordable public transport and smart land-use and distribution of (affordable and quality) services are all critical steps towards a more equitable, inclusive and just city. To intervene intelligently and advocate for change, however, we need to be able to use data to *see* these inter-relations more clearly and develop more holistic data-driven policies and projects.

Figure 18: Accessibility across Modes and Residential Developments in Nairobi (Campbell et al. 2019)

2.4. Better Understanding of Minibus Operations

Projects: Accramobility

Accramobility, a collaboration between the French Development Agency, Accra Metropolitan Assembly, Concordia University and Transitec, mapped out routes, stops as well as passengers' boarding and alighting of the city's minibuses (trotros) which gives further insights into the spatial distribution of the demand for transportation in Accra (Saddier et al. 2016, See Figure 18).

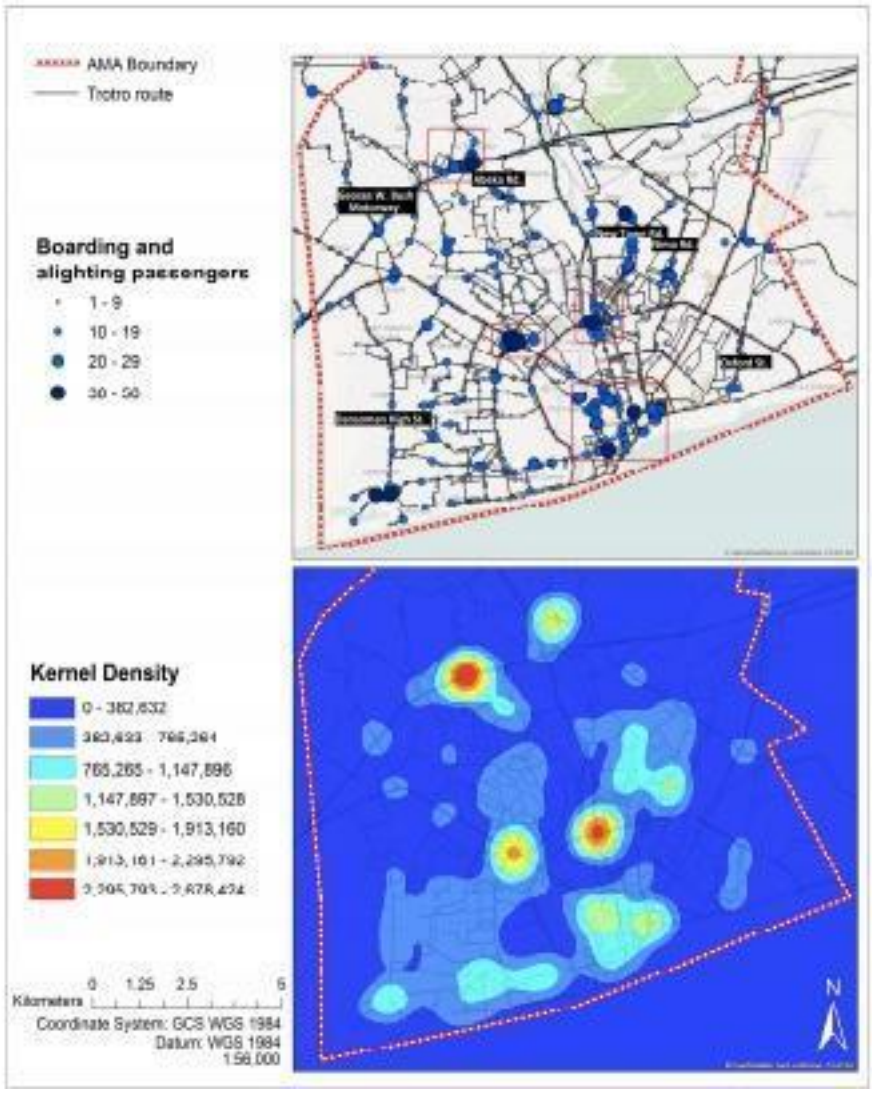


Figure 19: Spatial Distribution of Passenger Throughput (Saddier et. al 2016)

The aim of this work was to inform improved planning and interventions to upgrade operations. This work showed a relative stability of routes in Accra (Saddier et al. 2017) and also serious inefficiencies in operations.

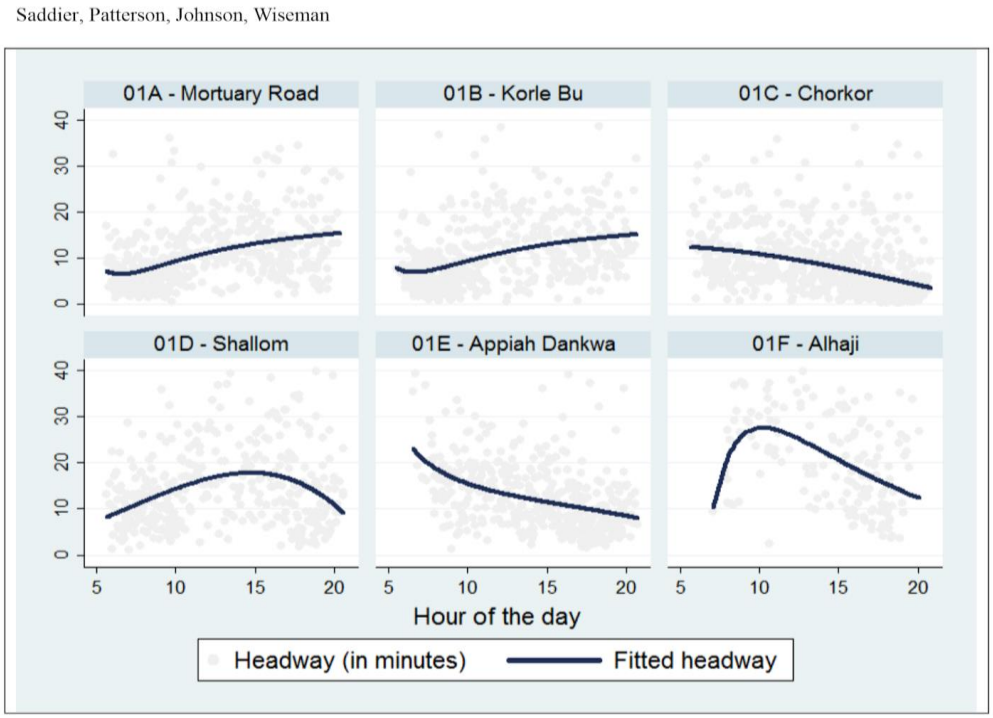
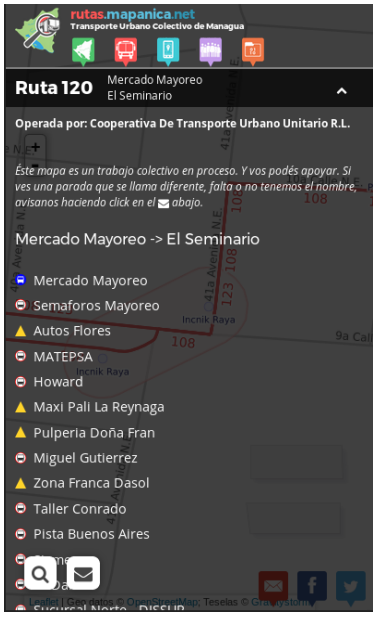


Figure 20: Departure profile for six different bus lines in Accra, Ghana (Saddier et al. 2018)

For example, vehicles appear to spend more time queuing than driving and do only a limited number of rotations a day suggesting wasted capacity (Saddier et al. 2018). Figure 19 shows the wide variation in headways for different minibus lines.

2.5. Formalizing the city with citizen participation: Popular definition of bus stops

Projects: MapaNica.net



Managua never had formal names for bus stops. The Regulatory Institute for Public Transport had handled the official bus stops with address descriptions and no name. In addition, Managua's public transport network consists of numerous unofficial stops that people know and bus drivers respect. The MapaNica.net bus mapping project conducted surveys with the citizens living close to bus stops in order to get an understanding of popular naming of each of the stops. This information was processed and consolidated to come up with a plan of unique names for all bus stops in the city. A reporting functionality allows citizens to suggest corrections through the project's website. Compare in Figure 21 the different icons for official (red circle) and unofficial (yellow triangle) stops and how they are presented on the website.

Figure 21: Stops view of Bussi in Managua

Interestingly in another case of mapping in Maputo, city officials improved and formalized some of the bus stops and a company called Ubi developed a transit screen as an experiment (Klopp and Cavoli 2019 See Figure 21 below).

Figure 22: Transit Screen in Maputo Source: Joaquin Romero.

2.6. Collaboration between and beyond the pilot projects

Projects: DigitalMatatus, MapaNica.net, Accramobility.

All three projects reveal the great potential that arises out of cross-city collaboration. Such collaboration is the building blocks of a growing global community of practice around open data creation and sharing in cities that currently have little to no data.

It is worth mentioning a few specific ways that these efforts are cross fertilizing and spreading ideas and tools for transit.

Collaboration between the projects:

1. Schematic map design: The schematic paper map design that came out of the DigitalMatatus created one of the first visualizations of paratransit systems. This achievement inspired projects in Managua and Accra to design similar schematic maps adapted to the local reality and requirements. Later such efforts would spread including to the Middle East (including Cairo and Amman).

2. Free and Open Source Software development: The MapaNica.net and Accramobility project collaborated together on various public Open Source Software. The project in Accra picked the web application *Bussi* developed by the MapaNica.net initiative. The teams collaborated on converting Open StreetMap to GTFS formats (*osm2gtfs)* opening the way to an integration of the data in the *Transportr* application through the *public-transport-enabler*.

3. Continental communities of practice: Two major groups emerged out of different initiatives one to promote Digital Transport for Africa (DigitalTransport4Africa) and one around Digital Transport for Latin America (DATUM) drawing in wide networks of collaboration to scale up this work.

3. Conclusions: Creating a Digital Commons for Inclusive Urban Transport Globally

Accumulating research and a growing number of mapping projects, including the ones profiled here, clearly demonstrate that it is possible to create valuable transport data in standardized form for all the world's cities and, by making it open, generate value and benefits across the globe. How to maximize these impacts remains to be further studied along with how this data can help improve monitoring and evaluation.

Many actors are learning and starting transit data work across the globe building emerging communities of practices around a digital commons approach to data and its use for transport. In addition, best practices are emerging[10]. However, these efforts remain sporadic and episodic, because of limited knowledge on how to use digital and mobile methodologies and tools in the service of the transport sector. In addition, despite increasing support and funding for sustainable urban public transport, financial and technical support for this work is . Yet this support is critical to institutionalize data creation, processing, publishing and updating as a fundamental part of capacity building for managing and planning for transport in most of the world's cities. How to finance and keep data fresh and most useful thus remains an unanswered question and may be approached in different ways and with different institutional approaches depending on city. Overall, the value generated from this data work in the transport sector is poorly understood, a gap this paper aims to address.

This overview also reveals that, as efforts to fill in the data gap grow, a need exists to develop principles and guidelines around transport data creation. New institutional configurations need to be developed to support cities so they can better understand, adapt, and maximally benefit from this digital transformation to provide better public transport. One key way to improve capacity, promote cross city learning and go to scale with data creation efforts is to promote collaborative open data networks to widely share data, learning and open source tools. This is also critical for sustaining data updates by building a skilled local community of users and creators-a wider local ecosystem.

In conclusion, the world's rapidly growing cities have a clear opportunity to leverage technology to build critical data infrastructure for varied forms of public and popular transport. They can use

[10] See the summary of best practices from mobilitydata.org.

this data to better visualize and integrate transport across modes as well as redesign networks and service for better access and in doing so build the blocks towards the option of a future mobility as a service platform. New technology and data tools can be applied in many creative ways to allow citizens, transport operators and government to see their whole system together, discuss it in new ways and reimagine policy interventions and projects and as a result move us closer to meeting our local and global goals.

4. Recommendations

1. Mainstream a “digital commons approach” to standardized open transport data creation in all transportation infrastructure support and technical assistance following best global practices and principles.

2. Support local data capacity building in cities and local ecosystems, encouraging the connection between cities, universities and their technology sectors in support of developing passenger information systems, data driven planning and monitoring for SDG 11.2 as well as institutionalization of updating processes.

3. Help build data sovereignty and local innovation hubs in cities through the digital commons approach as opposed to dependency on consultants and companies-often foreign- for data needs. This will help spur local research and knowledge that in turn can help cities be the agents of their own digital infrastructure. Spillover benefits from more diverse innovation communities will then travel across cities in South-South, North-South and South-North directions.

4. Use and invest in Open Source tools as a public good and leverage knowledge transfer, communities of practice and a sustainable approach in each city in the spirit of SDG 17.6 and 17.7.

5. Support inclusive efforts to build GTFS-flex in a way that incorporates challenges from cities with high levels of on demand, flexible mobility like paratransit.

6. Support development of privacy laws and public discussion of the ethics of data collection in all cities.

7. Build open and collaborative online platforms and hold meetings to incubate and foster a “digital commons” community to scale up open standardized data building and sharing efforts.

All these efforts will reduce the cost and increase the efficiency of data collection including for the SDG monitoring and also help create a greater emphasis on transparency in reaching our common goals. These interventions will help build up data as a fundamental infrastructure to manage and navigate complex and integrated multi-modal transport systems as well as for “proactive data-driven transportation planning solutions” (NACTO 2017). Finally, building new data capabilities and a network of data oriented public transport advocates across the globe will help many cities realize the very real possibility of moving more rapidly towards into a transit-oriented, safer, cleaner and more productive future- and in the process address climate change and hence help our planet as well.

References

Abisla, Ricky. 2019. “Open Transit Data in India” Paper written for the New America India US Public Interest Technology Fellowship.

American Public Transport Association. 2015. “Public Transportation Embracing Open Data” Accessed August 28 2018: https://www.apta.com/resources/reportsandpublications/Documents/APTA-Embracing-Open-Data.pdf

Avner, Paolo; Lall, Somik. 2016. Matchmaking in Nairobi: The Role of Land Use. Policy Research Working Paper;No. 7904. World Bank. Washington, DC. https://openknowledge.worldbank.org/handle/10986/25803 License: CC BY 3.0 IGO.

Barcelo, Jaume, Masao Kuwahara and Marc Miska. 2010. “Traffic Data Collection and Its Standardization” in J. Barcelos and M. Kuwahara (eds.) Traffic Data Collection and Its Standardization International Series in Operations Research and Management Science 144.

Behrens, R. D. McCormick and D. Mfinanga. 2016. *Paratransit in African Cities*. New York. Routledge.

Campbell Kayleigh, James Rising, Jacqueline Klopp and Jacinta Mwikali. 2019. “Accessibility across modes and residential type in Nairobi” *Journal of Transport Geography*. Volume 74. January 2019: 77-90. https://www.sciencedirect.com/science/article/pii/S0966692317308244?dgcid=coauthor

Cerrillo-I-Martinez, A. 2012. Fundamental interests and open data for re-use. *International Journal of Law and Information Technology*, 20:3, 203-222.

Colpaert, Pieter et al. 2017. Open transport data for maximising reuse in multimodal route planners: a study in Flanders IET Intelligent Transport Systems(2017), 11 (7): 397.

Colpaert, P. & Rojas Meléndez, J.A. 2019. Open data and transportation. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), The state of open data: Histories and horizons (pp. 215–224). Cape Town and Ottawa: African Minds and International Development Research Centre. http://stateofopendata.od4d.ne

Dobush, Grace (2016). “Nicaraguans are using crowdsourcing technology to prove that a good map can change your life”. *Quartz Ideas*. ttps://qz.com/742755/nicaraguans-are-using-crowdsourcing-technology-to-prove-that-a-good-map-can-change-your-life/ (retrieved 24.03.2019)

Dubé, L.; Bourhis, A.; Jacob, R. (2005). "The impact of structuring characteristics on the launching of virtual communities of practice". *Journal of Organizational Change Management*. 18 (2): 145–166.

8 Principles of Open Data. 2007. https://public.resource.org/8_principles.html Accessed 06/06/16.

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.*

Ferris, B., Watkins, K., Borning, A. 2010. OneBusAway: results from providing real-time arrival information for public transit. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1–10.

Francisco S. 2016. “Communities of Practice: Using systems thinking to co-create a better world”, *Pilot Projects Design Collective LLC*

Garaldi Borges, L. Oranges Cezarino, T. Capellaro Ferreira, O. Travençolo Muniz Sala, D. Lehr Unglaub, A. C. Ferreira. 2017. “Student organizations and Communities of Practice: Actions for the 2030 Agenda for Sustainable Development”. *J. Elsevier. The international journal of management education*, ISSN 1472-8117, Vol. 15, No. 2, 2017

Goldsmith, Stephen. 2013. “Open Data’s Road to Better Transit” *Government Technology* Accessed at: http://www.govtech.com/transportation/Open-Datas-Road-to-Better-Transit.html (retrieved 24.03.2019).

Hern A. 2018. “Citymapper launches bus-taxi hybrid Smart Ride in London” *The Guardian.* Accessed at *https://www.theguardian.com/technology/2018/feb/21/citymapper-launches-bus-taxi-hybrid-smart-ride-london-transit-app* (retrieved 24. 03.2019).

Hogge, B. 2016. “Open Data’s Impact: Transport for London: Get Set, Go!” ODI Impact Study. Accessed at http://odimpact.org/files/case-studies-transport-for-london.pdf (retrieved 24. 03.2019).

ITF. 2018. “Technology and Public Transport” Accessed August 29 2018 https://www.itfcongress2018.org/en/your-congress-your-voice/technology-in-public-transport/#.W4hiGEeyC04.twitter

Jennings G. and Behrens R. 2017. *The case for investing in paratransit strategies and regulation and reform.* Volvo Research and Education Foundations. May 2017.

Kitchin, R. 2014. The Data Revolution: Big Data, Open Data, Data Infrastructures and their Consequences. London: Sage.

Klopp JM, Williams S. Waiganjo P, Orwa D and White A. 2015. “Leveraging Cellphones for Wayfinding and Journey Planning in Semi-formal Bus Systems: Lessons from Digital Matatus in Nairobi” in *Planning Support Systems and Smart Cities*,” (Springer) in conjunction with the conference Computers in Urban Planning and Urban Management April 2015 MIT.

Klopp, J M and Petretta D. 2017. “The urban sustainable development goal: Indicators, complexity and the politics of measuring cities” *Cities* 63: 92–97.

Klopp J M. and Cavoli C. 2018. “The Paratransit Puzzle: Master Planning for Transportation in Maputo and Nairobi” in
(Im)mobilities in the city – creating knowledge for planning cities in the Global South and postcolonial cities edited by Tanu Priya Uteng and Karen Lucas Routledge.

Klopp J M and Cavoli C. 2019. “‘Mapping Mass Mini-Bus Transit in Maputo and Nairobi: Engaging ‘Paratransit’ in Transportation Planning for African Cities” *Transport Reviews* https://www.tandfonline.com/doi/full/10.1080/01441647.2019.1598513

Krambeck, H. 2015. Presentation on “The general transit feed specification (gtfs) and implications for international development” World Bank and WRI -Transforming Transportation 2015. Accessed at May 2019 at:
https://www.slideshare.net/EMBARQNetwork/the-general-transit-feed-specification-gtfs-and-implications-for-international-development-holly-krambeck-world-bank-transforming-transportation-2015

Lee-Gosselin, M. and Buliung R. 2012. “The role of ICTs in the transformation of the experience of travel” *Transportation* 39:873-876.

McHugh, B. 2013. “Pioneering Open Data Standards: The GTFS Story” in *Beyond Transparency: Open Data and the Future of Civic Innovation.* Edited by Brett Goldstein and Lauren Dyson. San Francisco: Code for America Press.

Mitric, S. 2013. “Urban transport lending by the World Bank: the last decade” *Research in Transportation Economics* 40: 19-33Neumann, A. 2014. A paratransit-inspired evolutionary process for public transit network design. PhD Thesis. Technical University, Berlin.

National Association of City Transport Officials (NACTO). 2017. City Data Sharing Principles: Integrating New Technologies into City Streets. Accessed August 29 2018 at https://nacto.org/wp-content/uploads/2017/01/NACTO-Policy-Data-Sharing-Principles.pdf

Neuman, A., Roder, D, Roder and Joubert J.W. 2015. “Toward a simulation of minibuses in South Africa” Journal of Transport and Land Use. 8 (1): 137-154.

Neumann, A & Joubert, J W. 2016. The “Minibus” Contribution. In: Horni, A, Nagel, K and Axhausen, K W. (eds.) *The Multi-Agent Transport Simulation MATSim*, Pp. 111–114. London: Ubiquity Press. DOI: http:// dx.doi.org/10.5334/baw.17. License: CC-BY 4.0.

Nyabola, N. 2018. *Digital Democracy, Analogue Politics.* London: Zed Press.

O’Brian O. 2017 “Smart Mobility and Open Data: A Global and Personal Perspective” Accessed 28 August 2018 at https://www.itdp.org/2017/11/20/smart-mobility-open-data/ (retrieved 24.03.2019).

Open Data Institute. 2018. Definition of Open Data. Open Data Handbook. Accessed May 2019 at http://opendatahandbook.org/guide/en/what-is-open-data/

Open Data Watch. 2018. “The Data Value Chain: Moving from Production to Impact”. Open Data Watch website. Accessed at https://opendatawatch.com/wp-content/uploads/2018/03/Data_Value_Chain-WR-1803126.pdf (retrieved 24.03.2019).

Pessoa, L., Reed, L. Tzegaegbe, J., Wong J. & Yan, B. 2011. Enabling Transit Solutions: A Case for Open Data. *Georgia Institute of Technology*, Fall.

Principles for Digital Development. 2015. Website https://digitalprinciples.org (retrieved 5.3.2019)

Probst, G., Stefano Borzillo, S. 2008. "Why communities of practice succeed and why they fail". *European Management Journal (2008) 26*: 335-347. Elsevier.

Reynolds M., Blackmore C., Ison R., Shah R., and Wedlock E. 2018. "The Role of Systems Thinking in the Practice of Implementing Sustainable Development Goals" *Handbook of Sustainability Science and Research, World Sustainability Series. Springer International.*

Rojas FM. 2012. "Transit Transparency: Effective Disclosure through Open Data" Transparency Policy Project, Ash Center for Democratic Governance and Innovation. Taubman Center for State and Local Government, Harvard Kennedy School Cambridge Mass. Accessed August 2018 at http://reconnectingamerica.org/assets/Uploads/20120828FINALUTCTransitTransparency8-28-2012.pdf

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-foropen-transit-data/.

Saddier, S., Z. Patterson, A. Johnson and M. Chan. 2016 "Mapping the jitney network with 14 smartphones in Accra, Ghana: The Accramobile experiment" *Transportation Research Record: 15 Journal of the Transportation Research Board*, (2581) 113–122.

Saddier, S, Z. Patterson, A. Johnson and N. Wiseman. 2017. Fickle or Flexible? Assessing Paratransit Reliability with 2 Smartphones in Accra, Ghana. Paper presented at the Transportation Research Board 94rd Annual Meeting.

Saddier, S and A. Johnson. 2018. "Understanding the Operational Characteristics of Paratransit Services in Accra, Ghana, A Case Study" . Paper presented at the Transportation Research Board 95rd Annual Meeting.

Scassa, T. 2014. Public Transit Data Through an Intellectual Property Lens: Lessons About Open Data. *Fordham Urban Law Journal* 41, 1759-1810.

Scassa, T., & Diebel, A. 2016. "Open or Closed ? Open Licensing of Real-time Public Sector Transit Data" *Journal of Democracy* 8(2): 1-20.

Schweiger, Carol. Transit Cooperative Research Program, Transportation Research Board, National Academies of Sciences, Engineering and Medicine.2015. *Open Data: Challenges and Opportunities for Transit Agencies* Washington DC: National Academies Press.

Senge, P. M. 1990. "The Fifth Discipline: the Art and Practice of the Learning Organization." *Doubleday/Currency, Print*, New York.

Shaheen S, E. Martin, A. Cohen, Apoorva Musunuri and Abhinav Bhattacharyya 2016. "Mobile Apps and Transportation: A Review of Smartphone Apps and a Study of User Response to Multimodal Traveler Information" Report for California Department of Transport Accessed at https://escholarship.org/uc/item/6m332192

SLoCaT. 2016. "Sustainable Development Goals & Transport". Partnership on Sustainable Low Carbon Transport. SLoCaT website: http://www.slocat.net/sdgs-transport (Accessed 30.04.2019).

Stokes, A. 2019. Governance and Transportation in Nairobi, Kenya: Understanding How Policy, Planning, and Levels of Governance Alter Mobility through a Multi-Modal Network Analysis Unpublished Thesis, Faculty of Architecture and Planning, Columbia University.

Sustainable Mobility for All. 2017. Global Mobility Report 2017: Tracking Sector Performance. Washington DC.

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/ReProgramming-Mobility-Report.pdf . Accessed March 4, 2015.

Transit Center. 2018. The Data Transit Riders Want. New York: Transit Center. Accessed at http://transitcenter.org/publications/transit_data/

Transportation Research Board. 2003a. TCRP Report 92: Strategies for Improved Traveler Information, Vol. 1., Washington, DC.

United Nations. 2014. World urbanization prospects: The 2014 revision. Department of Economic and Social Affairs. Population Division Accessed at http://www.un.org/en/development/desa/publications/2014-revision-world-urbanization-prospects.html

United Nations Statistical Commission. 2019. Accessed May 2019 at https://unstats.un.org/sdgs/metadata/

Urama K., Ozor N. and Acheampong E. 2014. "Achieving Sustainable Development Goals (SDGs) Through Transformative Governance Practices and Vertical Alignment at the National and Subnational Levels in Africa" African Technology Policy Studies Network (ATPS). Regional Practice Paper 2014.

Valentino-DeVries J., Singer N, Keller M. and Krolik A. 2018. "Your Apps Know Where You Were Last Night, and They're Not Keeping It Secret" New York Times. 10 December 2018. Accessed May 2019 at https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html

Van Wee. Bert, Karst Geurs and Caspar Chorus. 2013. "Information, communication, travel behavior and accessibility" *The Journal of Transport and Land-use*. 6 (3): 1-16.

Walker, J. 2014. "The Evolution of Logic in the Developing World" Accessed at "https://humantransit.org/2014/02/the-evolution-of-logic-in-developing-world-transit.html

Wenger E. 1998. "Communities of Practice: Learning, Meaning, and Identity". *Cambridge University Press*

Wenger, E.; McDermott, R.; Snyder, W. M. (2002). "Cultivating Communities of Practice". *Harvard Business Press*; 1 edition. ISBN 978-1-57851-330-7.

Wenger-Trayner, Etienne and Beverly. 2015. "Introduction to communities of practice". Wenger and Trayner website. Accessed at: https://wenger-trayner.com/introduction-to-communities-of-practice/ (retrieved 24.30.2019)

Wilkinson, P. 2010. "The regulatory cycle stalled? An assessment of current institutional obstacles to regulatory reform in the provision of road-based public transport in Cape Town, Africa" *Research in Transportation Economics* 29: 387-394.

Williams S., Waiganjo P., White A. Orwa D. and Klopp JM. 2015. The Digital Matatu Project: Using Cellphones to Create Open Source Data for Nairobi's Semi-formal Bus System" *Journal for Transport Geography* 49: 39-51. http://www.sciencedirect.com/science/article/pii/S0966692315001878

Williams S, Marcello E., and Klopp J. 2014. “Open Source Nairobi: Creating a GIS Database for the City of Nairobi to Provide Equal Access to Information” *Annals of the Association of American Geographers* 104:1, 114-130.

Wong, J. 2013. “Leveraging the General Transit Feed Specification for Efficient Transit Analysis” *Transportation Research Record* Vol 2338, Issue 1, pp. 11 – 19.

World Bank Open Transport Team. “An Overview of Open Transport in East and Southeast Asia” Accessed

World Bank. 2016. *Kenya Urbanization Review*. World Bank, Washington, DC. https://openknowledge.worldbank.org/handle/10986/23753 License: CC BY 3.0 IGO.