



FLOWMINDER.ORG

Using Mobile Phone Data: Benefits, Challenges and Opportunities

08 February 2024

Flowminder Foundation
Thomas Smallwood, PhD
Data Scientist



Our services

What
we do



**Mobile Data
Partnerships**



**Population
Distribution &
Mobility
Analysis**



**Geospatial Data
Analysis & Site
Placement
Optimisation**




**Capacity
Strengthening**

Our mobile data collaborations to date

Countries where Flowminder has collaborated with MNOs (present and past):

- Curacao (x 2 MNOs)
- Haiti
- Sierra Leone
- Ghana
- DRC (x 2 MNOs)
- Namibia
- Mozambique (x 3 MNOs via INCM)
- Nepal
- Papua New Guinea
- Senegal (advanced discussion)





To enable **decision makers** to access
the **data** they need to **transform** the
lives of vulnerable people, at scale.

Implementation & real impact requires collaboration

- The use of CDR data often involves many stakeholders e.g.:
 - MNOs
 - Government end users
 - Regulators
 - IGOs
- **Ghana is a good model of government partnership**
 - Ghana Statistical Services
 - Vodafone Ghana
 - Flowminder

Working with governments is an essential criteria for generating impact.



FLOWMINDER.ORG



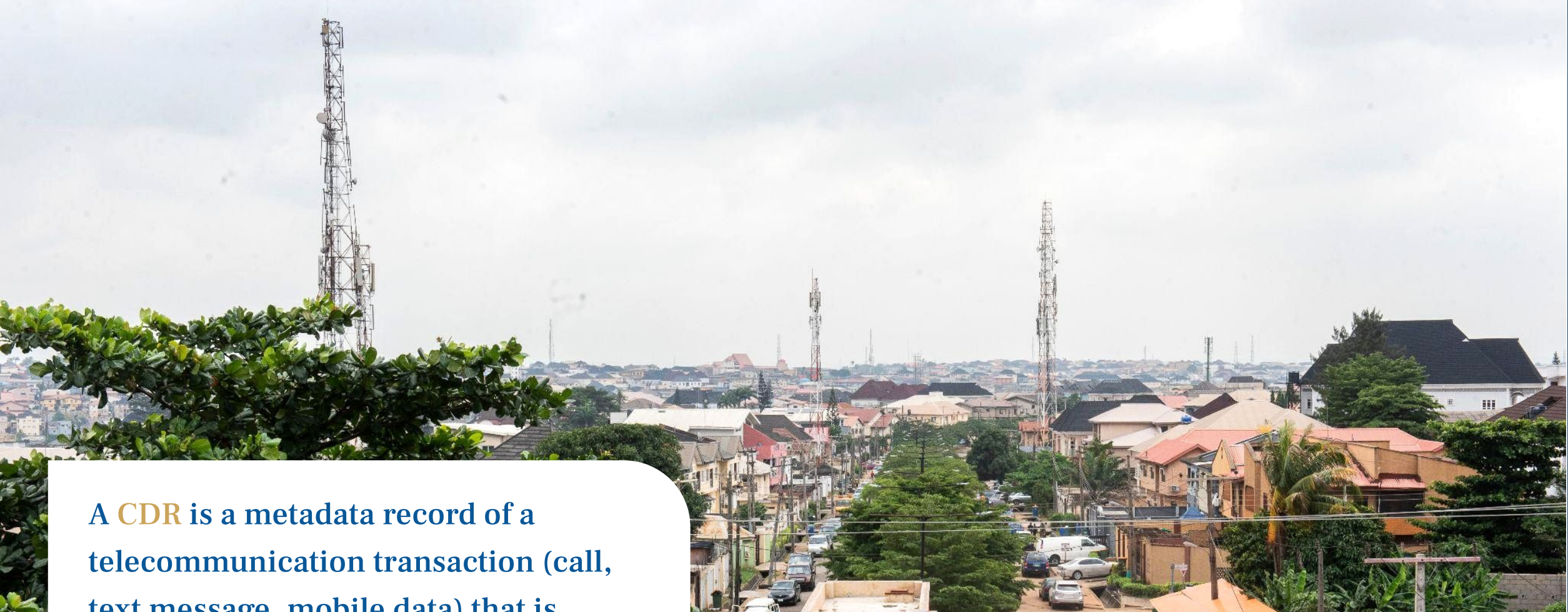
FLOWMINDER.ORG

“If we want to achieve the SDGs, we need to invest in data systems, and we are confident that our **successful partnership** with Vodafone Ghana & Flowminder is the way forward to **leverage data for good**.”

— Omar Seidu, Head of Demographic Statistics and SDGs Coordinator at [Ghana Statistical Service](#)



What the benefits and limitations of CDR data?

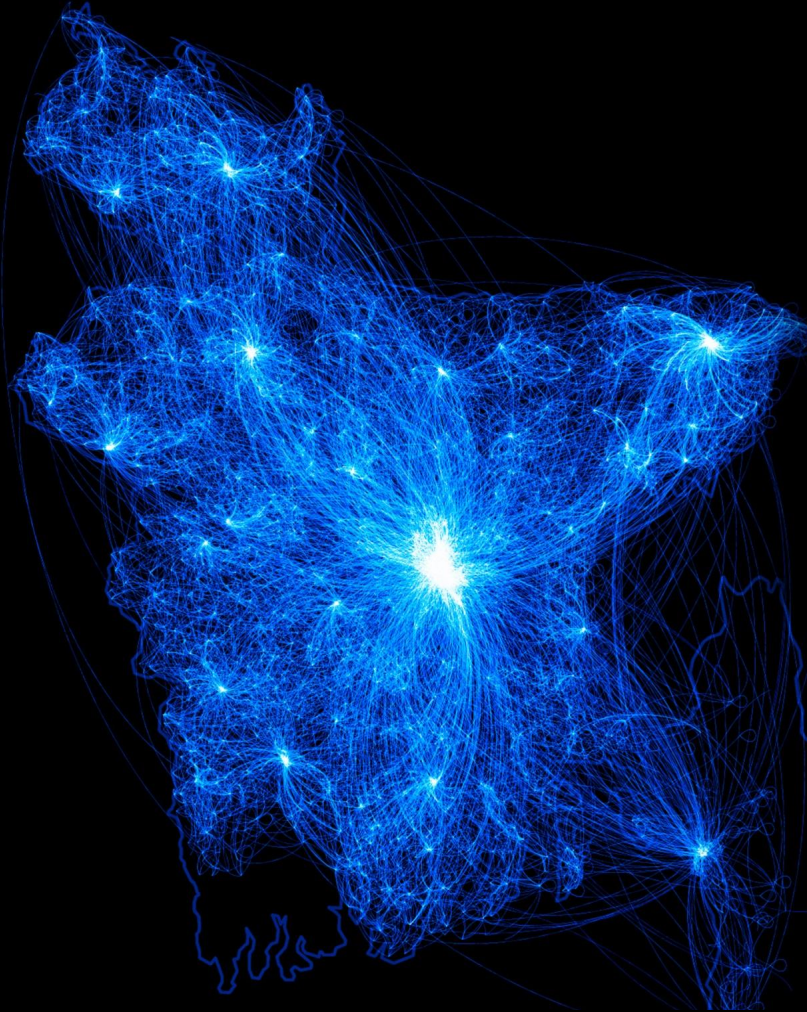


A **CDR** is a metadata record of a telecommunication transaction (call, text message, mobile data) that is stored by an MNO for billing purposes. It includes (as a minimum) a **subscriber ID**, a **timestamp** and a **cell tower ID**.

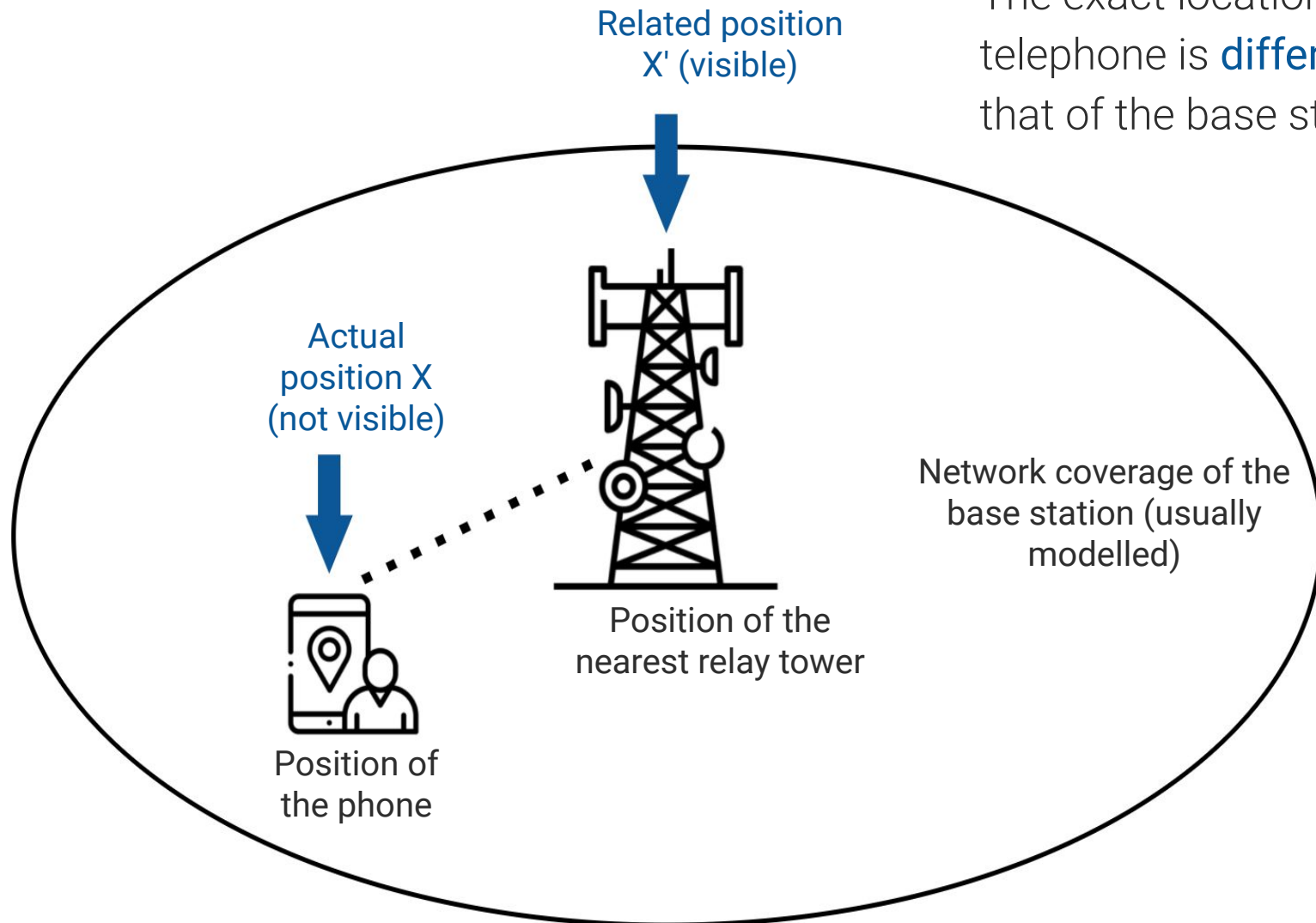
Cell Tower Data is a set of geographic data indicating the spatial position of each cell tower in a network. It includes (as a minimum) all cell tower identifiers, together with their **associated locations (typically latitude and longitude)**.

Benefits of using CDRs

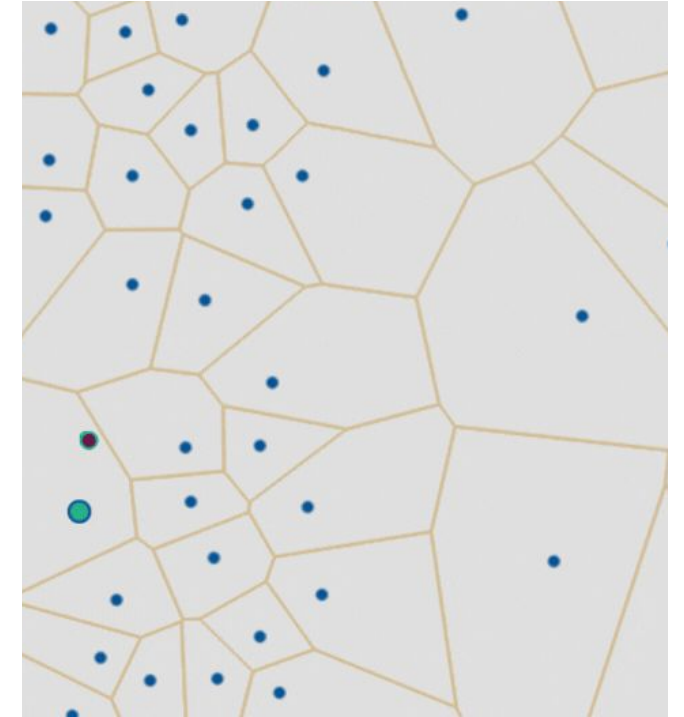
- The data are automatically generated.
- The dataset contains **billions of data points from millions of people** – large geographic and time scales covered.
- There is a continuous stream of data – **near real-time**.



Cell towers & subscriber location



The exact location of the telephone is **different** from that of the base station.

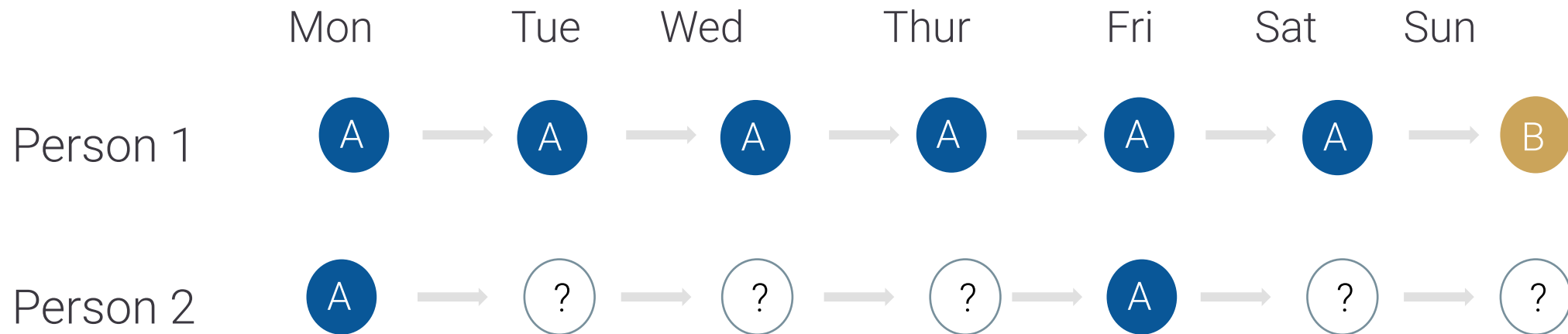


Application of CDR data in **geographic space for a user**. Each point represents a cell tower, each arrow the inferred movement of the user.

Temporal resolution

We only 'see' a subscriber when they use their phone.

If they don't use their phone on a particular day, we can't confidently say where they are on that day.



Ensuring privacy and transparency

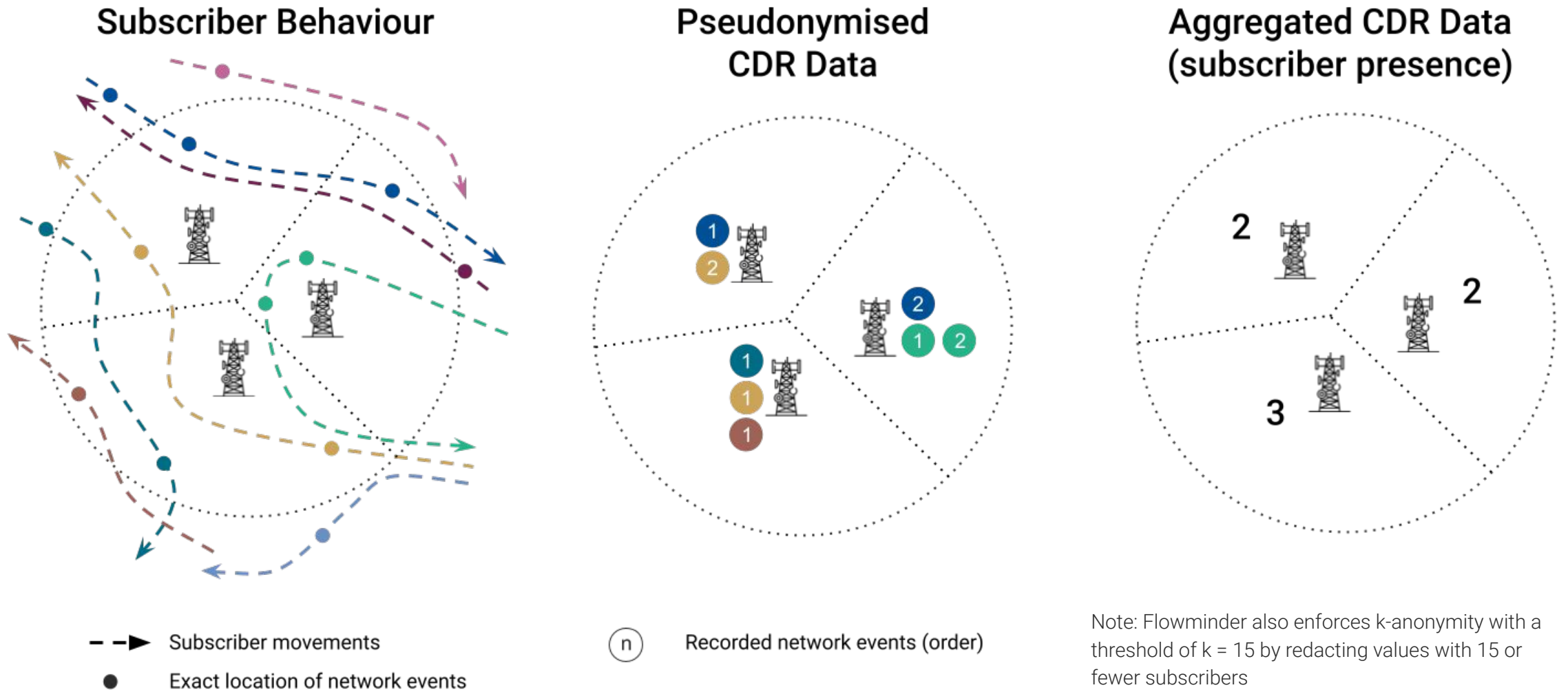
Key principles

- GDPR compliance throughout
- Transparency and peer review:
 - Detailed and open method descriptions
 - Open algorithms
 - Publications in peer-reviewed academic journals



CDR-derived
insights should
never permit the
identification of
individual
subscribers

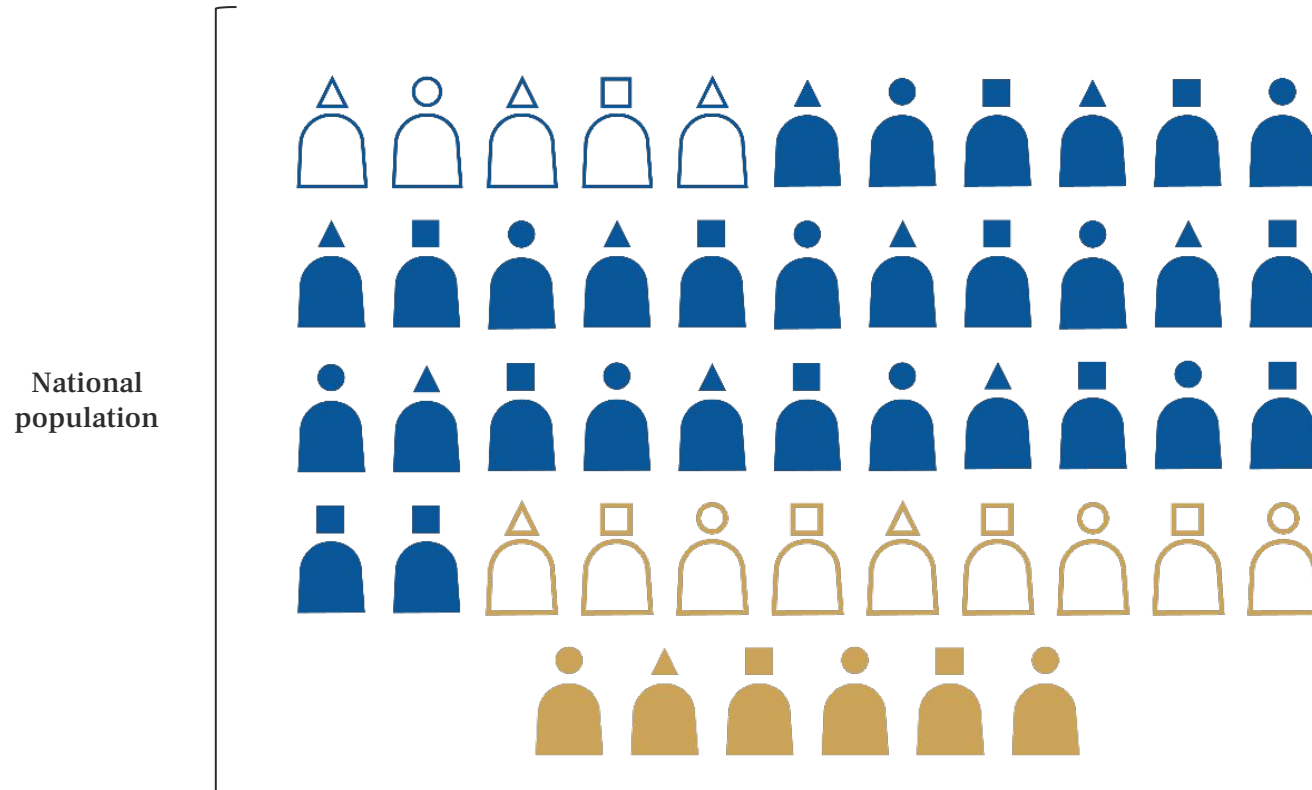
CDRs are aggregated to preserve the individual privacy of subscribers



Note: Flowminder also enforces k-anonymity with a threshold of $k = 15$ by redacting values with 15 or fewer subscribers

MNO subscribers are
not a random sample of the population,
nor can be assumed to be.

Addressing potential sampling biases



- No mobile subscription
- Subscribers to other mobile network operators (MNOs)
- "Inactive" subscribers of participating MNO
- "Active" subscribers of participating MNO

Additional data on demographics, phone use and mobility from field and phone surveys can help to **address biases** and therefore to **get the most out of CDR data.**

Addressing potential sampling biases



Sample for analysis*

- No mobile subscription
- Subscribers to other mobile network operators (MNOs)
- "Inactive" subscribers of participating MNO
- **"Active" subscribers of participating MNO**

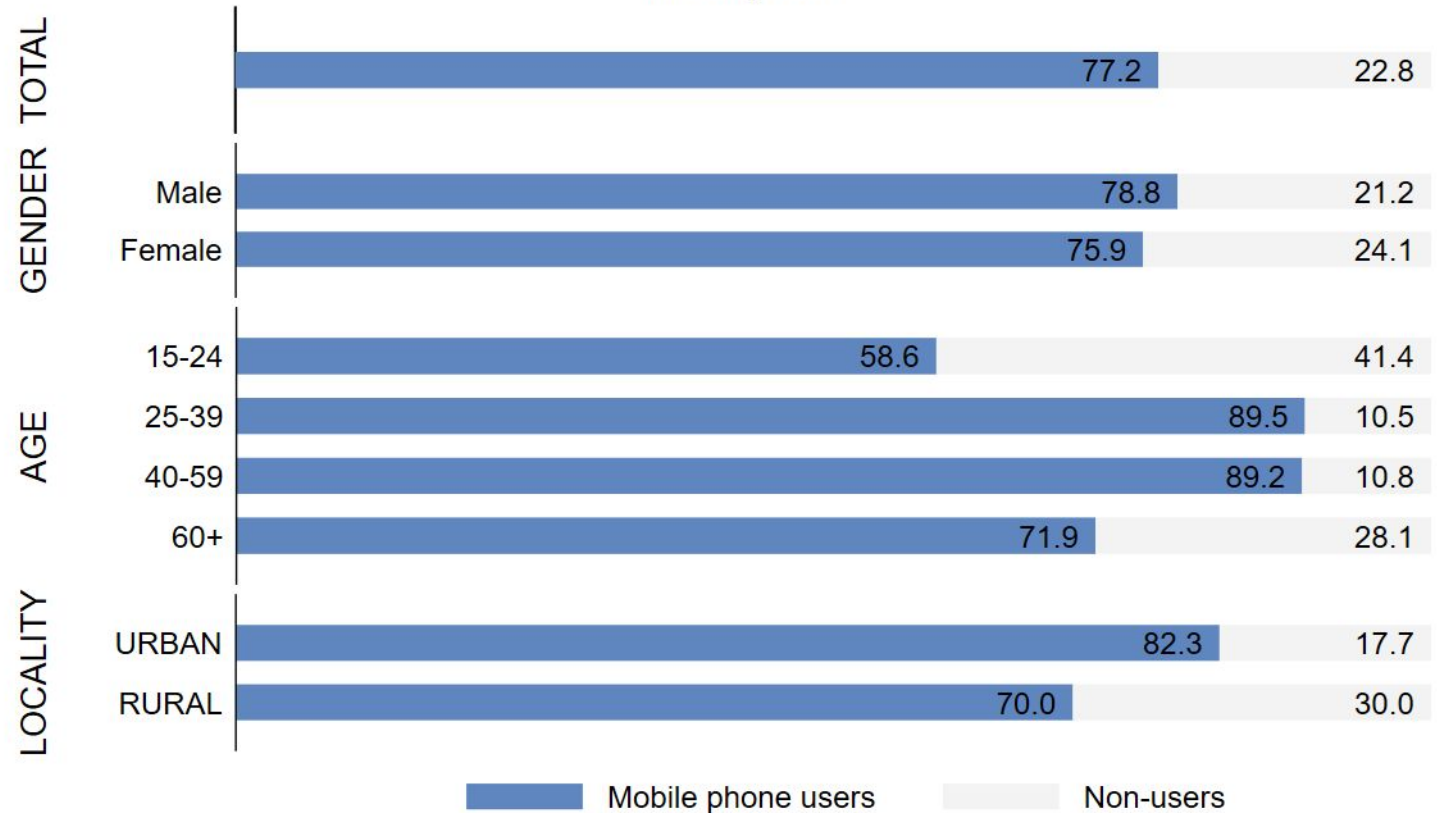
* Non-random 'sample' available for analysis

Additional data on demographics, phone use and mobility from field and phone surveys can help to **address biases** and therefore to **get the most out of CDR data.**

Mobile phone users by demographics | Ghana

- **Mobile phone use** is
 - More common among **men** than among women
 - More common among persons of **working-age** than among the elderly, and among children
 - More widespread in **urban areas** than in rural areas

Ghana: percentage of mobile phone users (ages 15+)
n=31,284

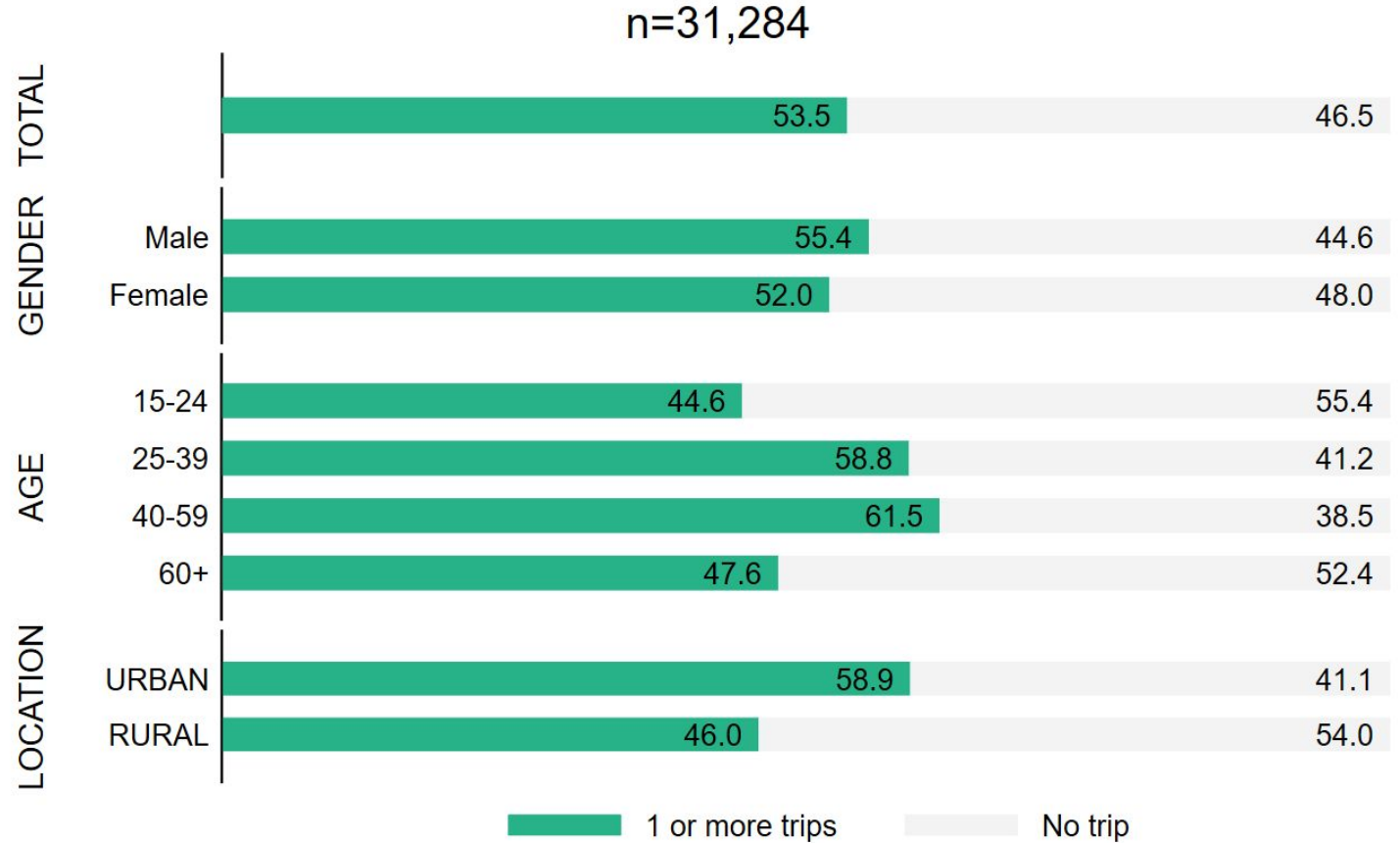


Source: AHIES 2022, Q3, weighted(pop_weight)

Mobility by demographics | Ghana

- Also, **mobility** is
 - More common among **men** than among women
 - More common among persons of **working-age** than among the elderly, and among children
 - More widespread in **urban areas** than in rural areas

Ghana: percent of population w recent trips 5+km (ages 15+)

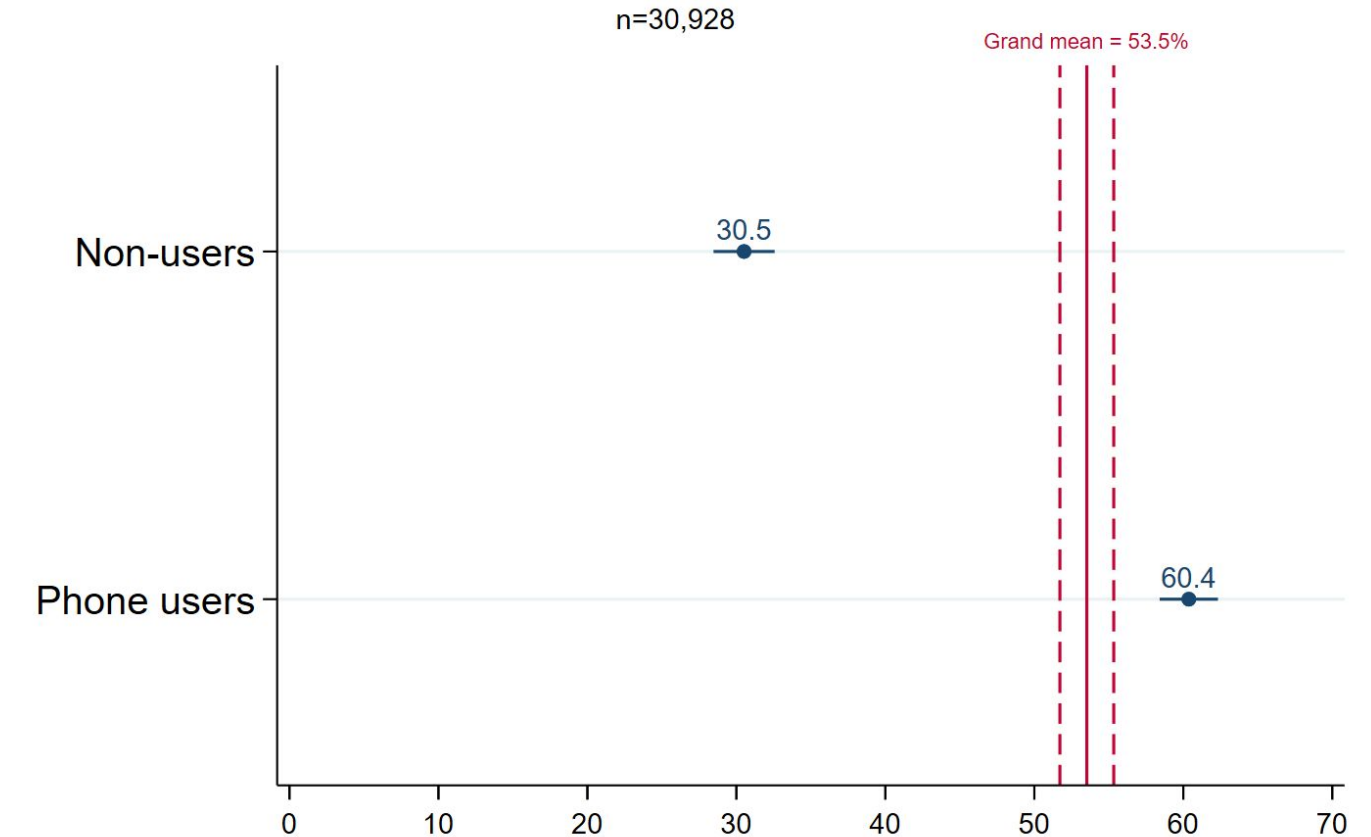


Source: AHIES 2022, Q3, weighted(pop_weight)

Biases due to mobility differentials | Ghana

- As a consequence, **mobility is often different** between mobile phone users and **non-users**
 - With mobile phone users showing **higher mobility**, on average, **than non-users**
 - Mobility estimates based on phone user data alone can **overestimate mobility** (can also be an underestimate for some types of mobility)
 - Not all mobility indicators, however, show such large differences between these two groups

Ghana: % of population who did a 5+km trip in past 3 months



Source: AHIES 2022, Q3, weighted (pop_weight)

Flowminder has recently developed estimation methods to **adjust for representation biases** & provide **population-scaled** estimates for

- **Relocations** from sub-region to sub-region, per month
- **Residents** per sub-region, per month

Bias-adjusted and population-scaled estimates

These estimates are based on

- CDR aggregates
- Primary & secondary survey data
- Existing population estimates



Population & Mobility Analysis



Applications



Official statistics

Estimating routine population
dynamics & migration

Routine mobility | Ghana

Data for Good partnership

Ghana | Routine Mobility Analysis for 2020-2022

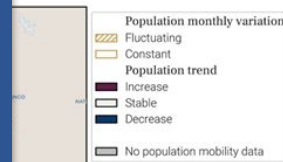
Insights into routine long term and short term mobility in Ghana

Beta version

Residents

... the estimated variation in the number of residents.

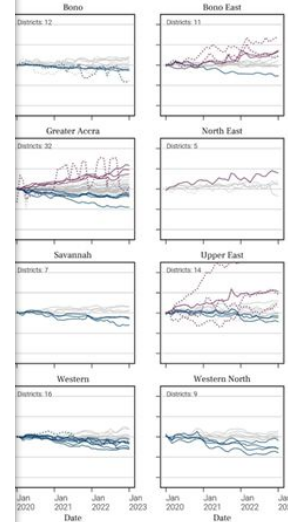
... or decreasing (and whether or not this trend changes compared to the usual trend) in the



...er the period 2020 to 2022.

Residents

Any trend (Constant) Any trend (Fluctuating)



Residents count, organised per region for each

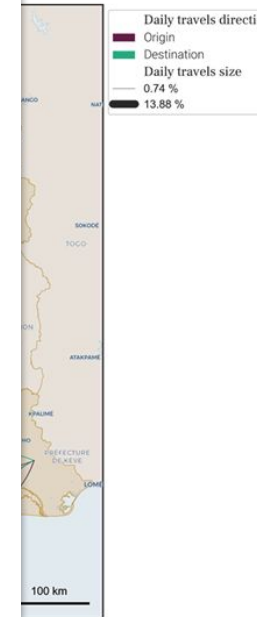
Population start (% of total)	Total population end	Population end (% of total)
14.7 %	4'275'058	14.0 %
3.4 %	819'639	2.7 %
62.6 %	19'181'744	62.8 %
3.1 %	951'785	3.1 %
11.2 %	3'659'086	12.0 %
5.0 %	1'668'352	5.5 %

...ciated districts and population statistics in

nts, i.e. the daily flows of people who are

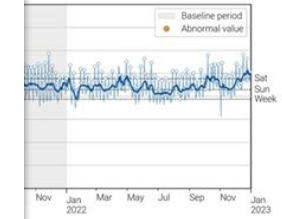
...rtion compared to all other flows) for the ...ted in the period 2020-2022. We then ...metres) for all travellers. Finally, we look ...est, and their associated daily time series

National level



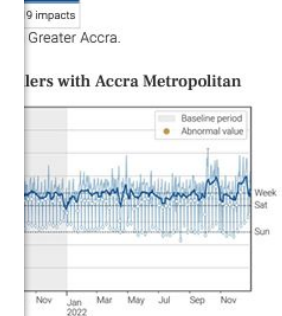
...ons in Ghana by relative size (compared 2022.

Residents with Greater Accra



...to and from Greater Accra with other ...values are categorised into weekdays, ...riod for each group.

Residents with Accra Metropolitan



...and from Accra Metropolitan with other ...values are categorised into weekdays, ...riod for each group.

Residents with Accra Metropolitan

Change (%)
9 impacts
-41.0
54.3

...Accra Metropolitan.

Applications

Disaster management

Displacement | Flooding on the Lower Volta River, Ghana

Report #1

Most recent data: 22 October 2023
Published: 07 November 2023

Ghana | October 2023 Lower Volta River Floods

FLOWMINDER.ORG

Long- and medium-distance displaced subscribers from flood-affected areas of the Lower Volta River: Situation as of 22 October 2023

Executive summary

The release of excess water from behind Akosomobo Dam on 11 October 2023, following heavy and unpredictable rainfall, has resulted in the flooding on communities in nine districts along the Lower Volta River. Large displacements have been reported as a result of the flooding (over 26,000, source: EU ERCC).

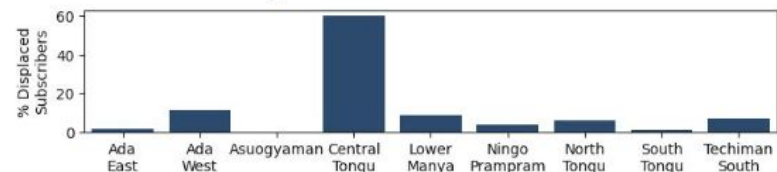
In this report, we focus on medium- (5–25km) and long-distance displacements (greater than 25km). We estimate that 1,430 Vodafone Ghana subscribers have been displaced from the areas around cell clusters in the affected districts. This represents only a portion of the total number of displaced persons, as these figures do not include those without a subscription to Vodafone Ghana or those who were displaced very short distances (less than 2km). We observe very few long distance displacements or displacements to urban centres, such as the Accra Metro area. The largest displacements observed were from the area around Mepe and Aveyime in North Tongu, to Mafi-Adidome in Central Tongu. We estimate that approximately 853 additional subscribers are residing in the area around Mafi-Adidome. We also observe that the increase in the number of subscribers displaced from the affected area has slowed. However, we do not yet observe any indications that subscribers are returning to their previous home locations.

Where have residents of the affected area been displaced to around Ghana?

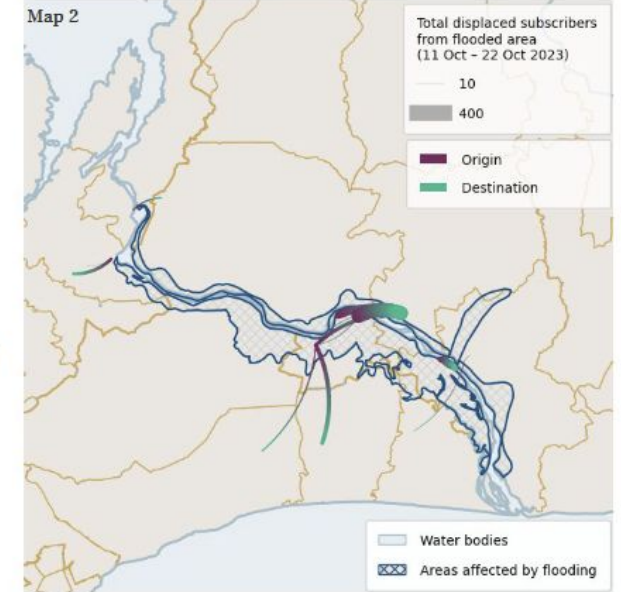
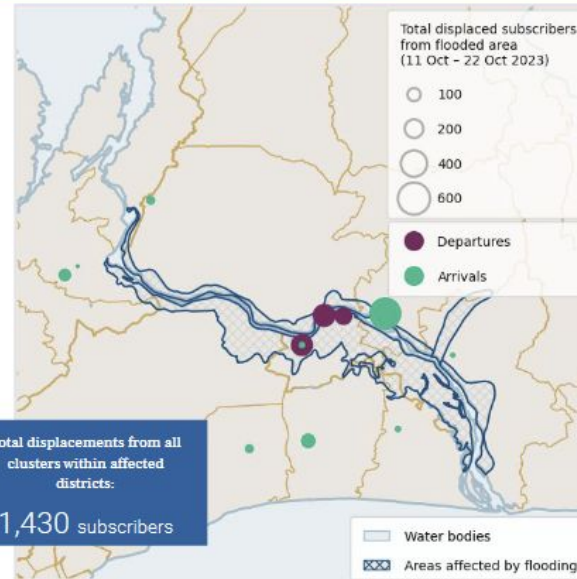
Origins and destinations of displaced subscribers

The bubbles on Map 1 (right) correspond to the number of displaced subscribers who have departed areas affected by flooding (purple) and the areas which have received displaced subscribers (green). The density of cell clusters means that we do not observe very short-distance displacements (less than 2km). The largest number of displaced subscribers depart from the areas around Mepe and Aveyime in North Tongu. The cluster with the largest number of arriving displaced subscribers is close to Mafi-Adidome in Central Tongu. We observed very few displaced subscribers outside of Eastern, Volta or Great Accra regions.

Graph 1 (below) shows that the majority of displaced subscribers (60.1%) relocated from the affected area to Central Tongu.



Graph 1: Percentage of the total displaced subscribers from the affected area who have relocated to each district



Conclusions

Based on anonymous aggregated mobile phone data from Vodafone Ghana, most displacements caused by the flooding on the Lower Volta River may have been local (less than 25km). Furthermore, most displaced subscribers are likely to remain within the affected or neighbouring districts.

The number of new displacements may also have slowed since the initial large displacement. However, we do not yet observe the population around the cell clusters starting to return to their pre-flood baseline, indicating that subscribers remain displaced.

Contact us: info@flowminder.org @Flowminder

The mobile operator data provided to Flowminder for this report was fully anonymised. No personal data was made available at any time (e.g. an individual's identity, demographics, location, contacts or movements). All results are aggregated (for example, subscriber density in a given municipality), meaning that they do not contain any information about individual subscribers.

FLOWMINDER.ORG

The mobile operator data provided to Flowminder for this report was fully anonymised. No personal data was made available at any time (e.g. an individual's identity, demographics, location, contacts or movements). All results are aggregated (for example, subscriber density in a given municipality), meaning that they do not contain any information about individual subscribers.

FLOWMINDER.ORG

Displacement | Gang violence in Port-au-Prince, Haiti

Report #3

Most recent data: 04 September 2023 **Haiti | August-September 2023 Gang violence, Port-au-Prince**

FLOWMINDER.ORG

Published: 08 September 2023

Displaced subscribers from violence-affected neighbourhoods, Carrefour-Feuilles and Savanes Pistache: Situation as of 04 September 2023

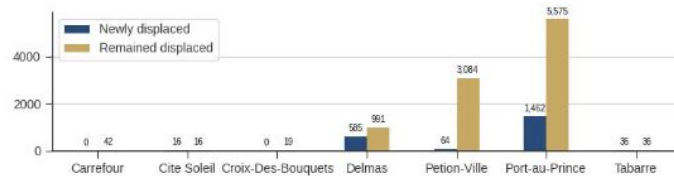
Executive summary:

Heightened gang violence since 12 August 2023 in the neighbourhoods of Carrefour-Feuilles and Savanes Pistache in Port-au-Prince continues to trigger significant population movements out of these areas.

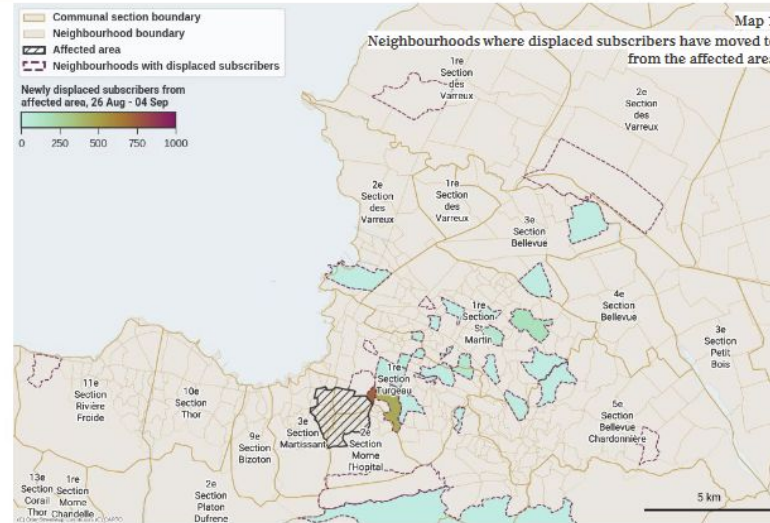
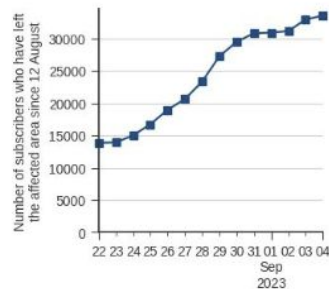
Based on anonymous aggregated mobile phone data, we estimate that up to 32,075 subscribers previously residing in the affected areas have since the start of the crisis left the area because of gang violence (most - if not all - are likely to be displaced and in vulnerable situations). Out of these, we can estimate that 9,770 subscribers are currently staying, primarily, in the neighbourhoods of Avenue Christophe (adjacent to Bas Peu de Chose) and Pacot in the Port-au-Prince commune, and Desire - Boutillier in the commune of Petion-Ville - see graph 1 and table 1), in line with IOM-DTM observations from 05 September. For the remaining 22,000 subscribers likely displaced, most are observed to remain within the Port-au-Prince commune, but in neighbourhoods we were unable to precisely determine (whether Avenue Christophe and Pacot and/or other neighbourhoods in the commune).

In the ten-day period from 26 August to 04 September, we estimate that 3,100 subscribers have been newly displaced, notably to Avenue Christophe and Pacot (commune of Port-au-Prince) and Siloe (commune of Delmas) (see table 1). For comparison, they were 2,000 between 19-26 August (8-day study) and 7,500 between 12-26 August (8-day study), showing that the situation is not yet stabilising.

Graph 1: Number of newly displaced subscribers (blue) and subscribers remaining displaced (gold) into neighbourhoods which we can precisely determine, summed per commune. We also observe that an additional 22,000 subscribers may be displaced, and most may be within the commune of Port-au-Prince.



Graph 2 | Total displaced subscribers: High bound estimate of the number of subscribers who have left the affected area since 12 August because of gang violence, and may still be displaced to date.



Total displaced subscribers since 12 August (high bound estimate)	Total displaced subscribers in known neighbourhoods since 12 August	Newly displaced subscribers out of the affected area between 26 August and 04 September
32,075	9,770	3,051

and newly displaced subscribers between 26 August and 04 September

Commune	Communal Section	Neighbourhood	Newly displaced subscribers this week	Total (Remaining) displaced
Delmas	1ère Saint martin	Peligre - Charbonnière	35	35 (*)
Petion-Ville	3ème Etang du Jong	Bremond	31	31 (*)
Delmas	1ère Saint martin	Jacquet Tybulle	27	27 (*)
Delmas	1ère Saint martin	Delmas 17	19	26
Delmas	1ère Saint martin	Delmas 75	21	21 (*)
Delmas	1ère Saint martin	Valme	20	20 (*)
Delmas	1ère Saint martin	Puits Blain	19	19 (*)
Tabarre	3ème Bellevue	Maia Gate - Clercine 8	19	19 (*)
Petion-Ville	5ème Bellevue Chardonnière	Pernier 24	-	17
Tabarre	3ème Bellevue	Galette Goureau	17	17 (*)
Petion-Ville	5ème Bellevue Chardonnière	Puits Blain	17	17 (*)
Port-au-Prince	1re Section Turgeau	Bourdon	16	16 (*)
Cite Soleil	1ère Section des Varreux	Hasco - Village Democratie	16	16 (*)
Port-au-Prince	1re Section Turgeau	Babiole - Debussy - Haut Turgeau	16	16 (*)
Croix-Des-Bouquets	1ère des Varreux	Village La Renaissance - Les Orangers	-	< 15
Croix-Des-Bouquets	2ème des Varreux	Santo	-	< 15

ates that subscribers likely displaced in this neighbourhood all arrived recently n 26 August and 04 September)

Contact us: info@flowminder.org @Flowminder

*Map for illustration purposes only. The boundaries and names shown and the designations featured do not imply official endorsement or acceptance by Flowminder, its partners, or donors.

FLOWMINDER.ORG

Contact us: info@flowminder.org @flowminder

FLOWMINDER.ORG

FLOWMINDER.ORG

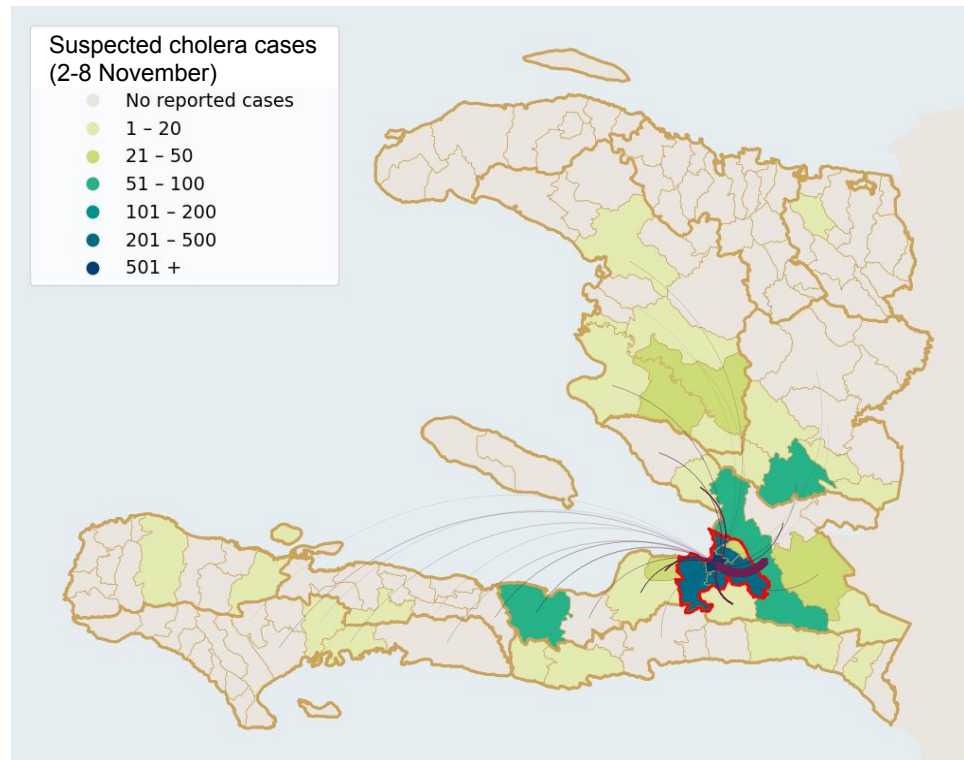
Applications

Public health & epidemiology

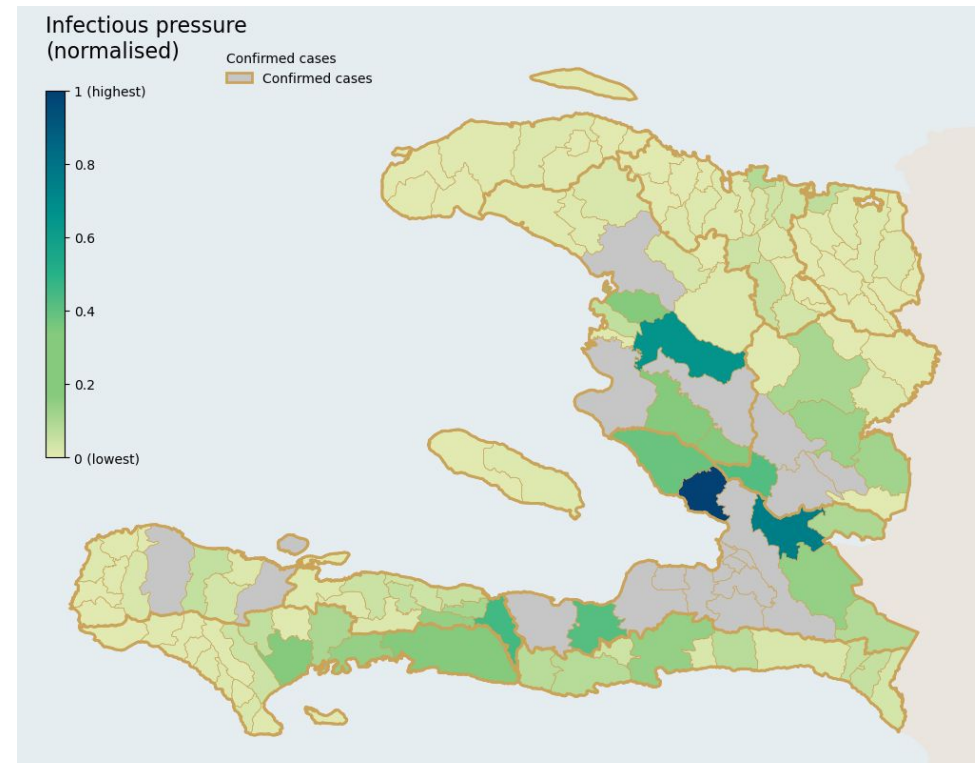


Modelling the spread of infectious disease | Cholera outbreak in Haiti

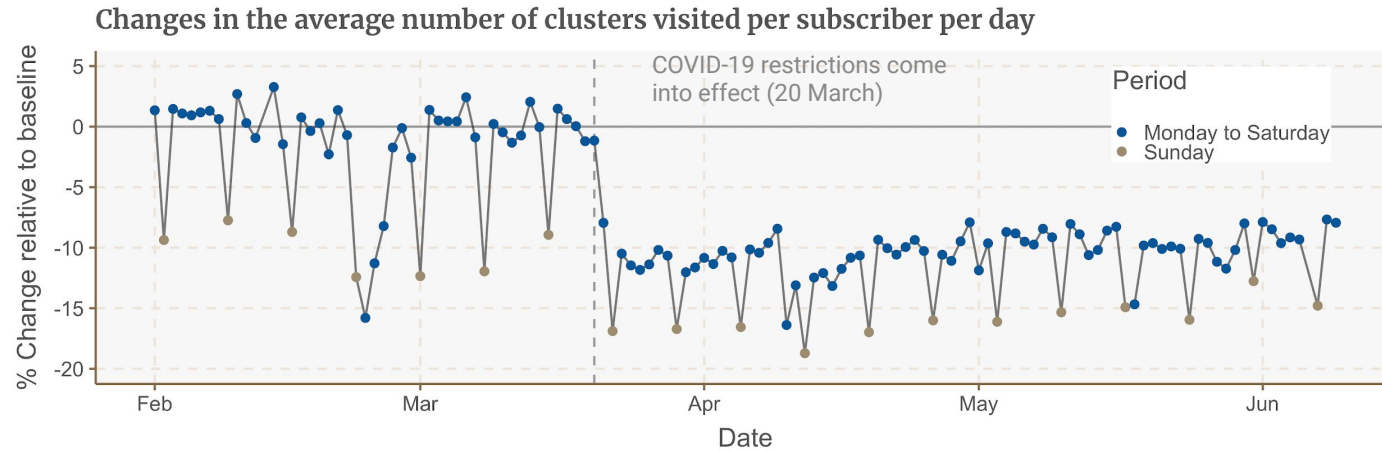
Flows from areas with high incidence in Port-au-Prince



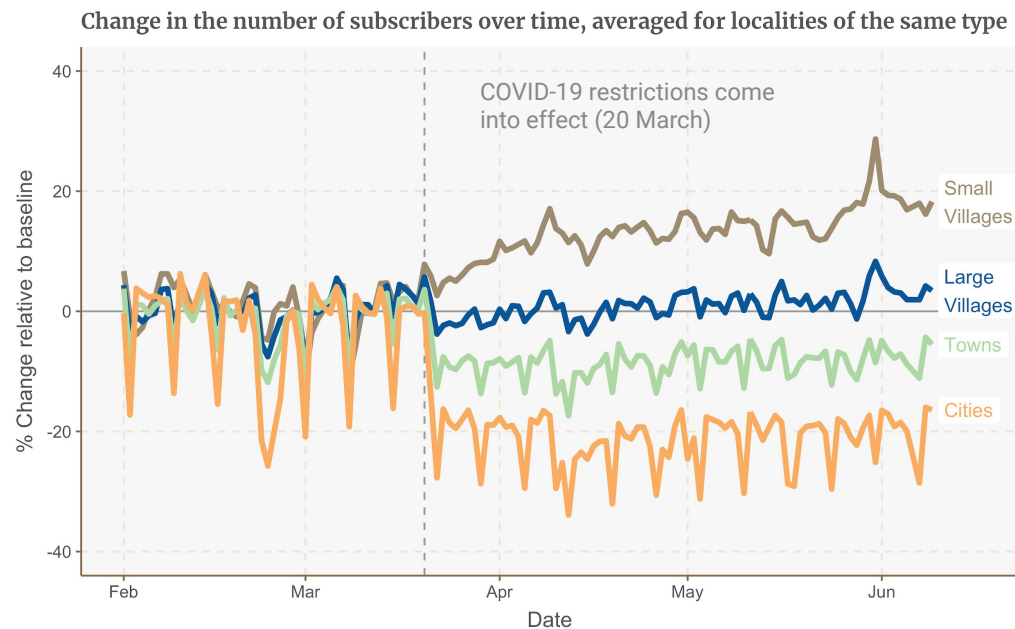
Infectious pressure on unaffected areas



Mobility Disruption | COVID-19 restrictions in Haiti



Mobility restrictions reduced the number of places visited during the day...



...but triggered a shift of population distribution from urban to rural areas (similar to changes occurring during the Christmas period)



Thank you | Q&As

Any
questions?

Learning Hub

UN Global Platform Learning Management System

Supporting a series of e-learning courses on various statistical and Big Data topics

Available through the UN Global Platform LMS
(learning.officialstatistics.org)



BADAN PUSAT STATISTIK

FLOWMINDER.ORG



FLOWMINDER.ORG

FlowGeek

- **FlowGeek is our open, online knowledge centre on CDR analytics**
 - created to **leverage the value** of CDR data and help **strengthen the community** of CDR data experts, enthusiasts and learners on the processing and analysis of such data.
- **Our content includes**
 - What are CDR data?
 - Applications of CDR data
 - Types of CDR indicators
 - Analyses and methodologies
 - Data governance and security



Each record contains the type of network event and the subscriber's identifier, as well as the time of the event and the cell site it was routed through.

The MNO stores these records in a database and may connect them with other information on the subscriber and their account, to inform their billing process. **These data can be analysed to better understand how people move within a country.**

Mobile phone metadata: CDRs

MNO	MNO_COUNTERPART	CELL_ID	REGION	EVENT_TYPE	TIMESTAMP
AA204Y1542DCA00	VENY782AD945G2E	451154211	north	voice	2016-10-10 15:35:25
AA204Y1542DCA01	ORBE728A008E51	451354312	north	voice	2016-10-10 20:03:45
AA204Y1542DCA02	ETB47088A0548C	451354312	north	voice	2016-10-10 21:21:56
AA204Y1542DCA03			north	voice	2016-10-10 21:59:32
AA204Y1542DCA04			central	voice	2016-10-10 22:42:23
845QW45CARVA5	RY2928A0V082930	476226941	south	ana	2016-10-10 08:13:21
845QW45CARVA6	ET0428CVAR036L	476126941	south	ana	2016-10-10 08:14:15
845QW45CARVA7	ET0428CVAR036L	476126941	south	ana	2016-10-10 08:14:59
845QW45CARVA8	RY2928A0V082930	476126941	south	ana	2016-10-10 12:41:01
845QW45CARVA9	RY2928A0V082930	476126941	south	ana	2016-10-10 13:10:45
845QW45CARVA10	RY2928A0V082930	476126941	south	ana	2016-10-10 19:20:43
845QW45CARVA11	RY2928A0V082930	476126941	south	ana	2016-10-10 18:09:32
845QW45CARVA12	RY2928A0V082930	476126941	south	ana	2016-10-10 18:09:32
845QW45CARVA13	RY2928A0V082930	476126941	south	ana	2016-10-10 18:09:32
845QW45CARVA14	RY2928A0V082930	476126941	south	ana	2016-10-10 21:59:32
CEW268W438EP2	ETB47088A0548C	451354312	north	voice	2016-10-10 09:01:10
CEW268W438EP3	ETB47088A0548C	451354312	north	voice	2016-10-10 21:59:20
CEW268W438EP4	ETB47088A0548C	451354312	north	voice	2016-10-10 21:59:20
DF33482W22Y7R	ORQ65088A0338Y	455193201	central	ana	2016-10-10 12:28:28

FLOWMINDER.ORG

Thomas Smallwood, PhD
Data Scientist, Flowminder

info@flowminder.org



www.flowminder.org



info@flowminder.org



[@Flowminder](https://twitter.com/Flowminder)