UN Mobile Data Training Workshop:
Obtaining access to mobile operator data

Tracey Li, PhD
tracey.li@flowminder.org
Discussion questions

1. What do you think are some of the challenges involved in NSOs gaining access to mobile operator data?

2. How do you think we can overcome these challenges?
Challenges

• Privacy concerns
• Ethical concerns
• Legal and regulatory obstacles
• Low incentives for mobile operators to share their data.
Privacy concerns – challenges

Mobile operator data is ‘personal data’ because it contains the location history of each person.

What concerns do you think people have about their data being shared?
Privacy concerns – solutions

Public communication about the project, from the start, is essential.

The aims of the project, and how the data will be used, should be made clear.
Being transparent about the processing and analysis methods that you will use will help to build trust.

- Publish methods in open-access journals and reports.
- Make code open-source.
Ethical concerns

What happens if the data falls into the hands of undesirable actors?

How do we prevent the misuse of data? Whose responsibility is this?

How do the risks compare to the benefits, and the risks of non-use?
Personal data are subject to strict national and regional legislations.

These laws specify how personal data needs to be treated. The laws are designed to protect the privacy rights of people.

Interpreting these laws, and ensuring compliance with them, can be difficult.
Legal and regulatory requirements

Many countries have national data privacy laws, and national regulatory bodies.

Do you know if there is such a law in your country?

Is there a national regulatory body?
How do we persuade operators to share their data?

Mobile operators must invest time and resources to provide data access.
There are also risks associated with sharing this data:

- A data breach would have a very negative impact on their reputation
- Miscommunication about the project will have a very negative impact
- The data have high commercial value to them, and competitors can take advantage.

How do you think we can persuade mobile operators to share their data?
What are the benefits for the mobile operator?
How do we persuade operators to share their data?

Ensure that you have taken all possible IT security measures to reduce the risk of a data breach, or of the data being misused.

Offer possible benefits to operator:
- corporate social responsibility (CSR)
- financial incentive (pay them)
- help them with building analytic capacity
Data access models
As a NSO, you can either have access to individual-level data, or aggregated data.

Individual-level data are the records of each individual person.

Aggregated data only contains the number of people in each region, or making a particular journey.
Individual-level data or aggregated data

Individual-level data is personal data.
It is subject to data protection laws.
Extensive IT infrastructure and IT security is required to securely process the data.

Aggregated data is not subject to many restrictions.
The mobile operator, or a third party, will aggregate the data and then pass to the NSO.
Storage and processing is much simpler.
Individual-level data or aggregated data - discuss

What are the advantage and disadvantages of choosing access to individual-level data, or aggregated data?
If you want access to individual-level data, one of the best models is remote access. Individual-level data never leave the mobile operator premises. All storage and processing is done on-site.
Accessing aggregated data (1)

The mobile operator can compute aggregates and provide these to the NSO. You only get the aggregates that the operator has computed, and have little control over what methods are used.

Alternatively, a third-party (e.g. Flowminder, Positium) can perform the aggregation, according to custom-designed methods.
Another option is for the mobile operator to install third-party software that provides an interface between their data and the NSO e.g. FlowKit (more tomorrow).

The NSO then has more freedom over which aggregates are produced and how, even though they never see the individual-level data.

The operator retains full control over who can access what data.
Which would be your preferred option?

Which of the options is most appealing to you? Why?

What difficulties do you think you will encounter if you proceed with this option?
Discussion: setting up a mobile data project
## Country statuses

<table>
<thead>
<tr>
<th>Interest</th>
<th>Strategy</th>
<th>Data access</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Egypt</td>
<td>Mongolia</td>
<td>The Gambia</td>
<td>Indonesia</td>
</tr>
<tr>
<td>Rwanda</td>
<td>Morocco</td>
<td>Kenya</td>
<td></td>
</tr>
<tr>
<td>Botswana</td>
<td>Jordan</td>
<td>Ghana</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Liberia</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tunisia</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>