Reviewer: 1

Comments to the Author
Great paper. A few comments:

One of the biggest deterrents of Data for Development initiatives is engaging with mobile network operators (MNOs) and creating incentives for MNOs to share data/analytics. It would be good for the sake of knowledge sharing that the authors mention the incentives or modalities on how data was shared. Was this a commercial agreement/social good/co-shared value proposition model of sharing data?

Added in 1.2 and 2.2.2

Another interesting aspect of Big Data projects is the model for data access. It would be good if the authors also highlighted whether Kinshasa Digital analysed data within the Telco environment or off site; the same with Flowminder. This is important to share as such papers offer insight to other data for development practitioners on modalities of how such work is done. And that such detail are often the difficult pieces that other practitioners may learn from.

Added in 2.2.2

The work with Flowminder, was the analysis done using Flowkit? If that is the case, a few lines on what Flowkit is may be ideal. Reason being; Flowkit is an open source platform for analyzing MNO data for development. It is thus important to highlight such a perspective to enlighten other practitioners that such tools exist to do away with the barriers in deploying algorithms.

Regarding Flowminder data access and processing, I'm not 100% sure they have used Flowkit here, I believe they might have used something a bit more advanced there. In general, I'd rather not detailings Flowminder work as this will need to be reviewed and approved by them.

It would be good to capture a lesson on MNO engagement. What lessons can they share on how to close viable data sharing agreements with MNOs. This in my experience working in the area of Big data is among the top challenges. So we would like to learn what lessons to draw from their experience.

I've had some comments regarding MNOs in 3-4 of the learnings at the end

I see the authors mention that there are three more use cases to be deployed. This is great for continuity and sustainability.

One of the key issues with data for development initiatives is the legal and regulatory framework that guides data sharing/data use in a country or jurisdiction. It would be interesting to learn if there is a role that the National Regulatory authority played or whether there exist enabling data protection laws in DRC that guided this initiative or other safe guard rails.

There was a quick mention of the regulator endorsement in 1.2 and I've added two other mentions in the learnings. It is true it is usually a key issue but it wasn't really the case in DRC since the regulator endorsed the project from the start.

Reviewer: 2

Comments to the Author
The manuscript presents an interesting overview of lessons learnt in the framework of a Business-to-Government initiative aiming at harnessing the potential of insights extracted from Mobile Network Operators data to help fight COVID-19 in DRC. The document contributes to the objectives of the special issue to which the manuscript was submitted. Moreover, the work offers a discussion on the issues at the interface between data science and policy, in line with the type of "commentary" type of publication. Here below a set of specific comments to the authors.

The policy statement is probably too long and is incomplete in the proof made available for review.

Shortened

It is not fully clear if the exchange of data/insights is a data for good initiative or there has been financial assistance to MNOs. This should be clarified better in section 1.2, while the description of the financial support to the organisations involved in the initiative could probably be mentioned in the acknowledgment sections. Also section 2.1 could be summarised focussing on lessons learnt rather than listing the communication/exchange events.

Added in section 1.2 and 2.2.2

It would be useful to the reader to know a rough market share of Orange DRC, Africell and Vodacom to understand the representativity of the data used in the different applications. In more general terms, it is not fully clear what type of data were made available and if there has been any activity to assess their quality: was it Origin Destination Matrix aggregate data, tracks or simply insights provided directly by operators? What was the granularity, and the frequency? Have the data or insights of the three MNOs been combined in any of the studies?

Added a section in 2.2.2. We're not allowed to share market share figures though.

It also seems that epidemiological analysis was not performed in the framework of this initiative. It is difficult to implement epidemiological meta-population modelling (for example) if access is not provided to Origin Destination Matrices of movements between areas.

We have initiated an epidemiological model with our analyst and the WB, added a comment at the end of section 2.2.2.

There is no reference to ethical considerations in the commentary, although avoiding discrimination and respecting fundamental rights is in the GSMA COVID-19 privacy guidelines cited (by the way, the link to the guidelines could be added as footnote).

I've added some references to ethical considerations when we mention the GSMA COVID-19 privacy guidelines

There is also no reference to the fact that the data or insights are somehow made openly available also to researchers and to the public at large, or the access is restricted to policymakers. From the data availability statement the latter seems more plausible. However, it is clear that open data strategies would maximise the uptake and use of the data/insights, is there any future plan in this direction? Has this been discussed or proposed to MNOs?

I've added some extra explanations in the data availability statement. The codes are all open source but the data itself not. You can make a request to the mnos though. Open data strategy is not really an option for mnos in general given the sensitivity of the data.

There are a few typos, additional proof reading is recommended. Here below just a few of them:

- "ressource" – page 1
- "… of Congo" instead of "… of the Congo" – page 1
- "evidence-based policymaking" should probably be "evidence-informed policymaking" as evidence is one of the elements of policymaking together with e.g. values – page 2
- reference to a Task Force that was not previously introduced/specified – page 2
- "picture below" should refer to a numbered figure, i.e. Figure 4 – page 8
- the text in figure 4 cannot be read – page 9
- "…has increasingly gain in…" – page 11
- the e-health solutions could be summarised in a short paragraph rather in a list of bullet points – page 12

Fixed

Reviewer 3:

Quality of the paper and its suitability for publication In its current form, the manuscript has several weaknesses:

• A general impression is that the learnings presented are not contextualized and interpreted in a greater setting. In order to allow for transfer learning to other geographies, this should be required as a minimum.
• General proofreading of the manuscript is needed. Several places the text needs rewriting in order to clarify and communicate better, several figures & tables are in French, which is not appropriate for this journal.
Fixed
• Referencing is in general not up to academic standards: Remove overlapping referencing between footnotes, bibliography, and Section 5.4 Data availability statement; add proper citations where there are hyperlinks in the text, and review and replace bibliographic entries containing hyperlinks that reference non-existent content.
Hopefully all is fixed now

A major revision of the manuscript is needed before I can advise publication.

Suggestions for improvement The following suggestions will strengthen the contribution, improve the scientific quality of the manuscript, and clarify ambiguities and inaccuracies:
• The Policy Significance Statement is incomplete.
Minor changes applied. Please advise what is lacking exactly in this statement if still incomplete.

• Section 1.1
o Data from Worldometer needs to be refreshed.
Done
o Add proper citations for links: "latest health infrastructure data", "recent mapping", and "2016-2020 Plan National de Development Sanitaire".
Done
o The term "evidence-based policy making" is not self-explainable, and should at least be defined properly, and preferably exemplified through concrete examples.
Removed the term

• Section 1.2

o The title of this section is "How did the collaboration between the mobile industry and health authorities start?"; and this question is not answered in the text.
Additional text

o There is a long list of partners, at least 12 in number, and several obvious questions arise: How were all these partners managed? What are their different roles, role descriptions and responsibilities, and how did each partner contribute? When partners have overlapping roles, how was this managed?
I've added a figure to illustrate and a text to explain

• Section 1.3
o Add proper citation for link: "adheres to adequate data privacy standards".
Done

• Section 2
o What is the defining characteristic of an enabler?
Changed the term of enablers
o It is unclear how this project uncovered the stated, four enablers; and they come across as ad-hoc. o There are standard frameworks that could help contextualize the findings, e.g., DM-CRISP within data mining, and/or Cohen and Levinthal "absorptive capacity" framework.
The goal here was not to apply a standard framework but rather simply to group together some characteristics that we experienced helped preparing the government to use MBD
• Section 2.1
o General comment to the supportive quotes in Sections 2.1, 2.4, and 3.1: The quotes are made by one of the three authors of the manuscript, and hence comes across as unfortunate and self-promoting, and does not support the impartiality of the work.
I removed the individual name of the author of the quote to keep only Ministry of Health. Although Eric is one of the tree authors, his contribution to this paper was limited and as ANICIIS is the primary user of the dashboards, we believe it does make sense to keep the quotes here. I've added a quote from presidential taskforce and Secretariat Technique
• Section 2.2.2 o Provide proper citation for (Academy, n.d).
o The two MBD solutions developed in the project were based on Flux Vision & Flowminder solutions. How much adaptation of these previous solutions were needed in order to produce the MBD solutions?
FluxVision only provided the aggregate data. KD processed this data and built the dashboard from there. Hope it is clearer in the text now.
o Figure 2 and 3: Figures should be explained properly. The text accompanying the figures are in French, and in order to support the English readership, please, convert the text in the figures to English, or add proper descriptions/explanations supporting the figures.
Done
o Replace "The picture below described the …" with reference to Figure 4.
Done
o Figure 4 + accompanying description on page 9: The discussion comes across as a bit misleading when emphasizing the staged approach, because as explained, one can be left with the impression that you have to do phase 2 and 3 before you can do phase 4. That does not harmonize with my experience with these kinds of projects.
I've added some caveat in the text
• Section 2.3 o Add proper citation/reference to the GSMA User-Centered Design process and elaborate on this choice.

Added some explanation. It is not really a choice, it is simply a classic approach when designing products and tools. It is not GSMA process.

• Section 2.4 o I suggest that the manuscript is updated with the learnings from the workshops taken place, along with follow-up meetings. These learnings are essential to understanding the landscape better.

o What is the purpose of Figure 5? This figure has the same language issues, as earlier figures (Figures 2 and 3).

Removed

o What is Bluesquare? I cannot see this partner listed in Section 1.2.

Bluesquare is mentioned briefly in the start

• Section 3.1 o What are "community-relays agents"? o What are the skills currently lacking, in order to carry out the additional list of use-cases?

Some explanation added

• Section 3.2 o Please rewrite the first sentence of this section and clarify the message/points being made.

Done

o Add proper citation for (Santé, n.d).

Removed

o What is the smartphone application "STOP coronavirus" doing?

Done

• Section 3.3 o Add proper citation for GSMA C19 privacy guidelines.

Done

• Section 4 o It is stated "…. to identify six initial learnings …. ". Note that the list contains seven items.

Done

o There lacks reflection and discussion of these seven learnings up against the four enablers mentioned earlier in the manuscript. In its present form, it comes across as there is no connection between the enablers and the learnings, and I challenge you on making this connection visible/clearer.

I've grouped the learning along the 4 driving forces

o Please, remove the quotes on words in this section: e.g. "value", "impact", "influence", "co-create". This is just confusing, and it is highly unclear what you are trying to communicate.

Done

• Section 5.4 o What are the restrictions relating to data availability?

Added

o The link https://github.com/djeffkanda/GSMA is not working

Fixed

o All links in this section should be converted into proper references.

Done

• Section 5.5 References o Replace (n.d.) with proper reference. o Correct misspelling of 'Dasdboard' – and this points to a page that requires login. o Dashboards, D (2020) needs to be a proper reference, and please note that this page also requires login. o IMF, I. M. references a file. This needs to be fixed.

Fixed

• Section 5.6 Annexes o The table in the annex is incomprehensible in its current form. I had to magnify the text 400% in order to learn that the whole table is written in French. Text needs to be converted into English, and layout of table needs to be changed in order to improve readability.

Removed