Catalyzing Innovation
ANNUAL WORLD BANK CONFERENCE ON LAND AND POVERTY
WASHINGTON DC, MARCH 25-29, 2019

DEMOCRATIZING THE DATA REVOLUTION: BRINGING LOCAL PERSPECTIVES TO THE SURFACE

LISETTE MEY, LAURA MEGGIOLARO

lisette.mey@landportal.org

Paper prepared for presentation at the
"2019 WORLD BANK CONFERENCE ON LAND AND POVERTY"
The World Bank - Washington DC, March 25-29, 2019

Abstract

With the inclusion of land indicators in the Sustainable Development Goals, the data revolution has very noticeably reached the land sector. New technologies to capture, monitor or analyze land data are increasingly being developed, for improved public service provision and beyond. These innovations could very well be the catalyzing factor that is necessary to bring this data where it can be put to good use to achieve land tenure security for all, at a speed and scale that would otherwise not be possible. However, one critical element of this data revolution is at risk to be overlooked: a multi-stakeholder, inclusive approach. The available data that is interoperable is largely from stakeholders from the global North. This paper highlights research into the interoperability-compliance of key land datasets in Africa and calls for a more democratized approach to the data revolution - ensuring local perspectives are not left behind.

Key Words: data ecosystem, open data, information landscape, land data, data gaps

1. Introduction

The "Data Revolution" is such a well-established and widely recognized concept within the sustainable development sphere now, that it is hard to believe it has only been an integral part of the development agenda for the past few years. With the development of the post-2015 agenda, the High Level Panel appointed by UN Secretary General Ban-Ki Moon expressed a need for such a 'data revolution' to enable the transformative action necessary to respond to the demands of an incredibly complex development agenda. The High Level Panel recognized that better data and statistics will help governments to track progress and make sure decisions are evidence-based, as well as strengthen accountability. Most importantly, the High Level Panel acknowledges the need for a multi-stakeholder approach to such a data revolution: "This is not just about governments. International agencies, CSOs and the private sector should be involved."

With the inclusion of several land indicators in the post-2015 agenda - the Sustainable Development Goals - the data revolution has very noticeably also reached the land sector. Data collection, analytics and monitoring have been hot topics of discussion within the land community. New technologies and innovations to capture, monitor or analyze land data are increasingly being developed and presented, for improved public service provision but also beyond that, almost offering to be a panacea to achieving land tenure security. These innovations could very well be the catalyzing factor that is necessary to bring this data to a level where it can be put to good use to instate the change we are all working towards, at a speed and at a scale that would otherwise not be possible.

However, one critical element of this data revolution as highlighted by the High Level Panel in 2015, is at risk to be overlooked: the multi-stakeholder, inclusive approach. In this paper, we aim to investigate whether the Data Revolution is a powerful catalyzer for positive change, or whether it is furthering the polarization and inequalities globally. We hope to trigger thoughts and discussion about how we can collectively ensure that this Data Revolution does not leave the

most important people behind: those whose lives we want to improve by securing their rights and access to land.

2. The Data Revolution: catalyzing change?

Having and using information has always been a powerful force for change, helping to fight corruption, enabling citizens to participate more fully in public life and allowing people from all walks of life to exercise their fundamental human rights. Considering data is now often referred to as the "lifeblood of decision-making"[1] or part of the "fabric of today's societies and economies"[2], it is hard to imagine that less than two decades ago, there was hardly any data flows happening online.

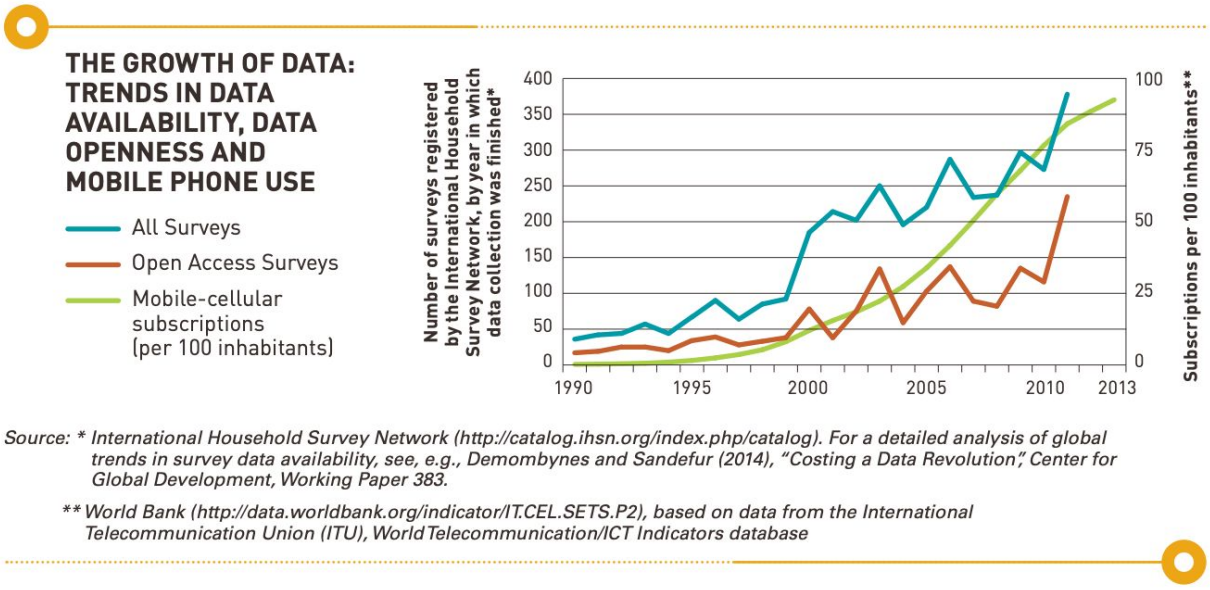

Figure 1: Exponential growth in availability & volumes of data[3]

[1] United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG), "*A World That Counts. Mobilising the Data Revolution for Sustainable Development*", November 2014.
[2] Calderon, A., "Bringing Power Into the Open", Open Data Charter: https://medium.com/@opendatacharter/bringing-power-into-the-open-966b271c41d2, February 2019.
[3] United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG), "*A World That Counts. Mobilising the Data Revolution for Sustainable Development*", November 2014.

In 2013, it was estimated that 90% of the data in the world was created in the last two years.[4] The graph above from the UN's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG) shows that not just the 'traditional' sources of data, such as the household surveys, have increased, but there has also been an exponential increase in new sources of data, such as mobile-cellular subscriptions.

This trend is also noticeable in the land sector. Where comparable land tenure data between countries was traditionally derived from National Agricultural Censuses[5], we have seen an exponential growth in apps and technologies to map plot boundaries and land tenure systems within communities. The theme of this conference alone is a testament to the fact that innovations to use and catalyze data and knowledge on land governance is a hot topic within the land community as well. The result of such innovations is that the number and types of data producers are increasing - anyone can be part of the data-discussions - and thus an enormous wealth of data is becoming available.

This continuously growing body of data as well as increasing demand for data is what we call the Data Revolution. These two elements however do not lead to the realization of the post-2015 agenda in and of themselves. The Data Revolution should be seen as a vehicle, a catalyzer that can enable this change to happen faster. Integrating different types of data, from different sources, and showing differences in numbers between traditional data sources with these new sources, enables the land community to get a more holistic view at the current situation, resulting in high quality *information* as opposed to raw *data*. More organized data that is machine readable and interoperable, allows data to move beyond numbers in spreadsheets used by data analysts and into applications and software that can be used by anyone.

[4] Petter Bae Brandtzæg, "*Big Data, for better or worse: 90% of world's data generated over last two years*" SINTEF, May 2013.
[5] FAO, "*World Programme for the Census of Agriculture*", http://www.fao.org/world-census-agriculture.

Thus, the Data Revolution has benefited the Sustainable Development Agenda, not only by creating a momentum for data comparison and bringing various rich sources of data together in order to increase its collective quality, but also by enabling this discourse to become more inclusive. Data is no longer confined to the conversations between data scientists, but is now brought into the open - resulting in a much greater degree of openness and transparency in the sector as a whole. With more people, with their varying perspectives on these issues, are entering into the discourse of the Data Revolution, data becomes more useful. Ultimately, the Data Revolution in its ideal form leads to more empowered people, better informed and therefore higher quality policies and decisions and greater participation from all parts of the population and increased accountability.

3. The Data Revolution: catalyzing polarization?

The Data Revolution in its ideal form brings different perspectives together and allows anyone and everyone to benefit from the richness of data available. However, we still have a long ways to go before the Data Revolution reaches its ideal form. We talked about the exponential growth in number of apps and technologies that can integrate interoperable data and thus make it more useful for a wider audience. This is a great development, but triggers an important question: what and whose data is suitable for such an application?

It is an illusion to think that if there is data and there is a certain technology or app, the two can automatically be combined and miraculously invoke positive impact at the local level. The word "data" is often loosely used to indicate any kind of information, knowledge or perspectives; this way, data can exist in many ways, shapes or forms. In order for "data" to be adapted and used in new technologies and apps however, it needs to be highly curated, adopted to international standards and published in a format, which altogether makes the data interoperable. This interoperability is exactly where the *ex*clusiveness of these technology and innovation trends lies. Because the datasets that are interoperable are those published by governments in the global North and large international agencies such as the World Bank and United Nations agencies. But

are grassroots or government data and perspectives from the Global South interoperable and therefore suitable to use in such technologies? According to our research, almost never.

The Land Portal has conducted a rigorous interoperability-compliance of datasets and databases that cover critical topics of land governance in four countries in Eastern Africa as well as South Africa. The critical topics have been identified through careful analysis of indicators in several global and regional monitoring initiatives, including the SDGs but also the Voluntary Guidelines on the Responsible Governance of Tenure, Monitoring & Evaluation of Land in Africa (MELA) and the Africa Data Revolution report, among others. An ambitious scoping study, involving desk top research, stakeholder surveys as well as peer review by multi-stakeholder platform, provides the basis of our data interoperability assessment. This assessment is based on internationally recognized criteria, such as those developed by the World Web Foundation, Open Knowledge International as well as Open Data Watch, the European Union and OECD. This research will be followed by capacity building efforts in-country, specifically targeted to the identified data holders and the gaps in interoperability-compliance in their own datasets, ultimately aimed at strengthening the local data ecosystem in the country.

At the time of writing this paper, the Land Portal has completed the interoperability-assessment of one of four pilot countries in East Africa. The assessment for Kenya has been completed and at the conference we will be able to present results for Tanzania, South Sudan and Uganda as well. The results for Kenya alone are staggering. Next to none of the local sources studied thus far comply with the internationally recognized principles of open data and interoperability. While at the global level, we keep talking of 'data', our research shows that 80% of key land information from local sources is actually to be found in *documents*. Actual raw data comes largely from international organizations, often from the United Nations or large research institutions, and the government. Though the latter category can hardly be called consistent in opening up its data, especially when it comes to land data.

Not only is there an inequality when it comes to data provision between global and local sources, but there are also significant gaps in access and use of data between users in the global North and the global South.[6] Many people, and a majority of those people are based in the global South, are excluded becoming part of the Data Revolution because of issues related to language, poverty, lack of capacity (in skills to use data but also lack of technological capacities to build and work with data infrastructures), remoteness, prejudice and discrimination. Take the example of Nicaragua, where the price of a mobile broadband subscription exceeds 10% of the average monthly GDP per capita, compared to France where this is only 0.1%.[7]

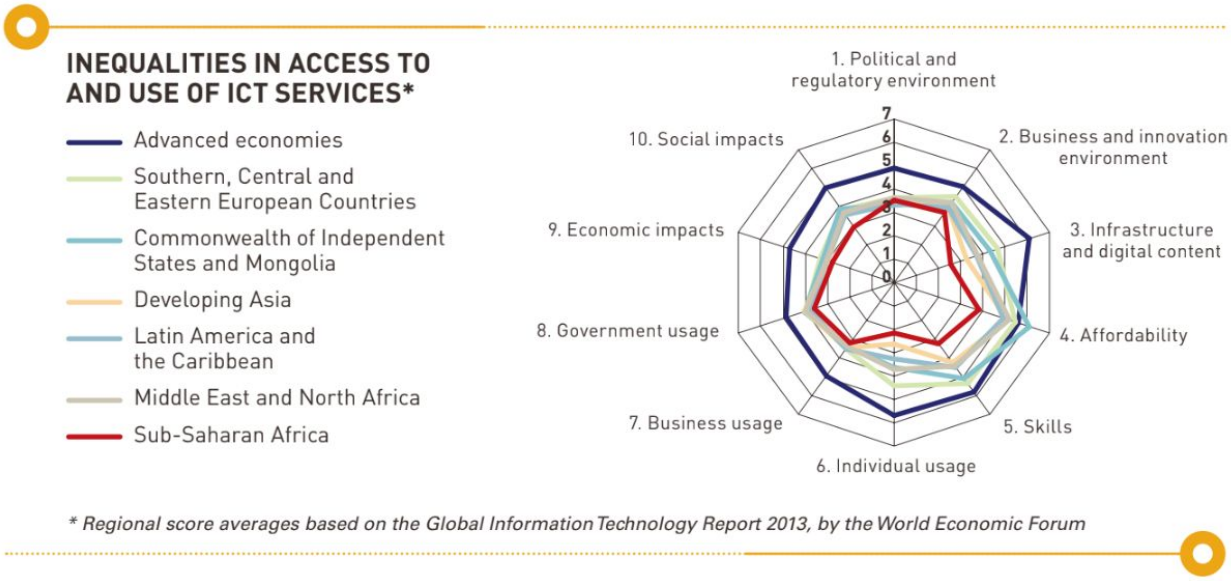

Figure 2: Inequalities in access and use of ICT Services across the world[8]

[6] United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG), "*A World That Counts. Mobilising the Data Revolution for Sustainable Development*", November 2014.
[7] ECLAC, "*Latin American Economic Outlook 2013: SME Policies for Structural Change*": http://www.cepal.org/publicaciones/xml/5/48385/leo2013_ing.pdf, 2014.
[8] United Nations Secretary-General's Independent Expert Advisory Group on a Data Revolution for Sustainable Development (IEAG), "*A World That Counts. Mobilising the Data Revolution for Sustainable Development*", November 2014.

4. The Data Revolution in the Land Community

Globally and across the entire development sector, we see the potential of the Data Revolution but recognize that we are not quite there yet. The Data Revolution poses several risks, including that of having a polarizing effect on inequalities, particularly between the global North and South, but also opportunities to bridge those gaps and enabling more inclusive discussions. What does this mean for the land sector?

We have seen how the amounts of innovations and applications to capture and collect one's own land data have increased enormously over the last years. The inclusion of land indicators in the Sustainable Development Goals has created an unprecedented momentum for data collection and monitoring within the land community. The Data Revolution has clearly extended to the land sector, yet renowned data initiatives such as the Open Data Barometer and the Open Data Index systematically rank land ownership data as the most closed and inaccessible dataset by governments worldwide. Unfortunately, this trend is not only occurring within government; the entire land sector is lagging behind in talking about and dealing with data.

The discussion around data in the land sector often does not focus on openness or interoperability, but a cry for high quality and up-to-date data. This cry is shared across all land stakeholders, though very likely the interpretation of what makes data of 'high quality' differs per stakeholder group. Doubt about reliability of data and distrust among actors are significant themes in the discourse on land data. Therefore the Data Revolution in the land sector probably is more behind than other communities within the development sector.

We need to overcome distrust between actors, competition over funding and fear of misuse of data. Sharing data and information should become a common practice by all stakeholders, accompanied by a standard set of metadata that includes the date of publication, the source of the data and methodologies. This way, when the data moves outside a particular database or website

and into applications or other repositories, users can still determine whether this data is of value for them or not.

Then intermediaries and other stakeholders have a responsibility to ensure that this data is not only valuable to specialized data analysts and scientists. People require certain skills and know-how to be able to digest the data and assess their trustworthiness or quality. Being able to assess quality and reliability of data is a level of data literacy that is, in our opinion, critical to achieve meaningful results in the sustainable development agenda and making the Data Revolution more inclusive.

While we believe that data analysis is an expert skill and should not be considered as something anyone can or should be able to do, we do not believe it is useful or effective to keep this type of discourse between experts exclusively. We may be skeptical of the notion that an inclusive data revolution entails training any- and everyone to be a skilled data analyst, but we are in complete agreement with the driving factor behind this notion: citizens and communities should be included and empowered to have a voice in the debate that ultimately aims to improve their daily lives and livelihoods.

Different methods can be applied to include citizens and communities in the debate. Not only should everyone be educated to be data literate to the extent where individuals become critical thinkers and learn to assess reliability of data, experts that possess the analytical skills have a societal responsibility to make data understandable to a wider, non-expert audience. Such efforts focused on data inclusion provide a basis for collective understanding, interpreting and managing data-driven decisions and discussions among all people.[9] Including and empowering (vulnerable) citizens and communities in such a way has the potential to increase their resilience in solving local problems and equips them with the necessary tools and skills to keep their governments accountable and transparent.

[9] Bhargava, R. et al, "*Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data*", Data-Pop Alliance, White Paper Series, MIT Center for Civic Media, United States of America, 2015.

The natural next question is, how do we make the data understandable for a non-expert audience? Bhargava et al phrase this in a very fitting way: "by making big data smaller".[10] "Small data" distinguishes itself from big data through one fundamental difference: unlike big data, the context about which the data tells a story plays a vital role.[11] It means bringing the big data back to a scale and in a format where more people are able to digest, understand and engage with it.

Using appropriate software tools to visualize data, for example, can be a highly effective way to communicate data and get the message across in a way that is understandable to the general public. Visual information is said to be processed 600,000 times faster than text and consumers of information are said to retain 80 percent of what they see, while only paying attention to 20 percent of what they read.[12] In a world bombarded with information, clear and concise visuals are what can grab a person's attention. In our case, without engaging visuals that depict the physical land we are talking about, we are only telling half of the story.

Democratizing the Data Revolution in the global ecosystem of land data requires efforts from all stakeholders in the land community. We all have a collective responsibility to share our knowledge, enable others to use it and to utilize our individual expertise in a way that brings value for our respective target audiences. It requires a shift in culture of dealing with data to look beyond our usual networks and collaborate with and involve other people.

5. Conclusion

The inclusion of several land-related indicators in the Sustainable Development Goals has brought the Data Revolution to the land sector and created an unprecedented momentum for land

[10] Bhargava, R. et al, "*Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data*", Data-Pop Alliance, White Paper Series, MIT Center for Civic Media, United States of America, 2015.
[11] Idem.
[12] Long, K., "*Infographic: Why visual content is better than text*", Ragan, United States of America. https://www.ragan.com/infographic-why-visual-content-is-better-than-text/, 2014.

data collection, analytics and monitoring. The Data Revolution brings an incredible opportunity to the sustainable development sector as a whole. It is the vehicle that, in its ideal form, leads to more empowered people, better informed and therefore higher quality policies and decisions and a greater participation from all parts of the population and increased accountability.

However, we still have a long ways to go before the Data Revolution reaches its ideal form. We talk about 'data' as if everybody has it and has the ability to use it, but there are clear gaps between those "who know" and those "who do not know". Particularly in the global South, perspectives are shared not in the form of raw data, but predominantly in the form of documents. Access to and capacities to use data are also significantly lower in the global South as opposed to the global North.

The reality is that in the global South, the "Data Revolution" debate focuses on the availability of data, the reliability of data and the question whether or not data is up to date. One can hardly be surprised that 'openness' of data or its 'interoperability' are not high on the priority list of data holders in the global South. However, this ongoing discourse provides us with an important opportunity to include data management, accessibility and interoperability into this debate in the global South. After all, openly accessible and interoperable data is - as mentioned - suitable to be put to great use through new technology and innovations. Immediate and large-scale proof of impact of the available data has proven to be a major incentive in the past for more actors to participate in the data revolution. Hopefully, this will lead to more actors willing to open up their data and share it in a way that it can also create impact we all seek to achieve.

There is a critical need to democratize the data revolution worldwide and not let the big international actors dominate the Data Revolution as they have done over the past years. Accuracy of data or reliability of it should be determined by bringing different perspectives together. If change is sought to be invoked at the local level, surely those perspectives are essential and we should not leave them behind.