

**12th World Telecommunication/ICT Indicators Symposium
(WTIS-14)**

Tbilisi, Georgia, 24-26 November 2014



Background document

**Document INF/3-E
19 November 2014**

English

SOURCE: Telefonica

TITLE: Big Data for Social Good: Opportunities and Challenges

Big Data for Social Good: Opportunities and Challenges

Nuria Oliver, PhD, nuria.oliver@telefonica.com

We live in a world of data, of big data. Our digital and physical interactions are increasingly leaving digital traces behind, which leads to a *big data revolution*.

Mobile phones are a powerful source of large-scale human behavioral data as it is the most widely adopted piece of technology in human history. In addition, mobile phones are personal devices that we carry with us all the time and are always connected, leaving a digital trace behind. Therefore, they can be seen as *sensors of large-scale human activity*.

Interestingly and very importantly, the analysis of mobile phone data can be done anonymously (*i.e.* without looking at any personal information) and in an aggregated fashion (*i.e.* always reporting aggregated results, never individually). For example, by only counting the total number of phone calls that each cell tower is handling or the number of phones that move between towers.

In my research team at Telefonica, we have been carrying our research in what we call Big Data for Social Good since 2008. The purpose of this research is to understand the value of large-scale human behavioral data (as it is captured passively by the mobile network infrastructure) for positive social impact.

Since then, we have been able to shed light on important questions such the impact that decisions made by governments have on the propagation of a pandemic [1,2,3,4,13,14], the automatic inference of socio-economic status in a developing nation [15] or the detection of floods from significant changes in the pattern of activity in cell towers [11]. These projects have arrived to important findings that are of great interest to governments and international health, development or emergency support organizations [1,2]

However, these projects are just a seed. For this seed to grow into something larger, a tree, there is a need to collaborate with the institutions that know the realities of the countries of study and have the decision making power to implement some of the findings. In addition, there should be a combination of mobile data with other types of data –such as transportation and traffic reports, weather and natural disaster information, emergency services locations.... – to fully reach the potential of Big Data for Social Good.

And the tree is starting to grow. We have already ongoing collaborations with United Nations Global Pulse, the World Food Project, the Mexican Government, the

DataPopAlliance, the GSMA and with MIT, among others. We are learning that some of the world's most important challenges can be understood and tackled through the analysis of big data. As Robert Kirkpatrick from UNGP states, "we should be able to work together, private and public sector, to find ways to harness big data for massive positive social impact, both safely and responsibly".

In addition to these focused collaborations, which are the branches of our tree, the ambitious goal is to plant a forest; a forest of big data for social good. Hence, some of this anonymized data is also opening up to a variety of organizations and research teams so that there is collaboration regarding this topic. For example, in September 2013, Telefónica carried out our first Datathon for Social Good at Campus Party (London) in collaboration with the Open Data Institute. Twelve teams participated in this datathon and analyzed both aggregated mobile data and open data and proposed a number of interesting projects, including a crime prediction model for London [10,5,6,8].

While the promise of using big data for social good is large, so are the challenges encountered to be able to realize this promise, as has become clear in light of the recent Ebola outbreak -the worst in our history, with almost 5,000 deaths and over 13,000 infections [17]. Unfortunately and despite the existence of large-scale human behavioral data that could be tremendously useful to help combat Ebola, we are only now - a few months after the outbreak -- starting to put into place coordinated efforts towards the analysis of mobile phone network data and what this tells about the spread of the disease.

Understandably, an important concern is the impact on privacy. The good news is that extensive research conducted by a range of academic teams demonstrates that it is possible to both analyze large-scale human behavioral data (*e.g.* levels of activity at the cell tower level, mobility patterns...) and preserve privacy. All data is typically anonymized using state-of-the-art encryption algorithms. In addition, data is usually analyzed in a highly secure and protected environment (*e.g.* the mobile operator premises) by authorized personnel. No analysis should be undertaken that would ever identify individuals. In addition, only the resulting aggregated, non-sensitive analyses (*e.g.* population mobility estimates, aggregate statistics...) would be made available to relevant aid or government agencies. Of course, an extremely clear code of conduct and ethical protocol would need to be followed by the authorized personnel handling anonymized individual data. A framework to share best practices among mobile phone operators and world recognized research teams would be extremely helpful in this regard.

A second challenge arises from the lack of expertise to be able to carry out such analyses in a technically sound way. However, technical difficulties should not be a barrier either, as there is a body of work illustrating how to carry out this type of analysis.

Moreover, there are highly skilled data scientists worldwide – including ourselves – and strong support from organisations, such as the ITU, ISOC, GSMA and United Nations Global Pulse – who are ready and willing to assist in the process, particularly to ensure that all data handling is carried out in an ethical and anonymous manner, always respecting local data privacy laws.

Finally, there are regulatory barriers and legal uncertainties that would need to be addressed. In particular, standards, practices and regulation would need to incorporate trust mechanisms for giving access to data for humanitarian purposes and in specific conditions (*e.g.* handling of the data by trusted partners, clear code of conduct and ethics when handling personal data, definition of where and how the data would be stored, guarantees that the data will be used only for humanitarian purposes, etc...) [16].

The potential to have positive impact and help save lives is immense. I truly hope that we can find a way to realize this the full potential of big data for social good. It's an opportunity that we cannot afford to miss.

References:

- [1][Using Big Data to Fight Pandemics](#) by Nuria Oliver. TechCrunch. Nov 2014
- [2]"[Ebola: Can big data analytics help contain its spread?](#)" by Matthew Wall. BBC News. Oct 2014
- [3]"[Nuria Oliver: What Big Data and the Mexican Pandemic Taught Us](#)". [WIRED](#). October 2013.
- [4]"[Combating Global Epidemics with Big Data](#)" Nuria Oliver. [The Guardian](#). September 2013.
- [5]"[Mobile Phone Data can help predict Crime one Month in Advance](#)" Motherboard. Sept 2014
- [6]"[Mobile Data could be Used to Help Predict Where Crime's About to Happen](#)" Gizmodo. Sept 2014
- [8]"[Researchers use cell-phone data, not precogs, to predict crime in London](#)" Engadget. Sept 2014
- [10]"[Once Upon a Crime: Towards Crime Prediction from Demographics and Mobile Data](#)" A. Bogomolov, B. Lepri, J. Staiano, **N. Oliver**, F. Pianesi, A. Pentland
[16th ACM International Conference on Multimodal Interaction \(ICMI 2014\)](#)
- [11]"[Flooding through the Lens of Mobile Phone Activity](#)"
Pastor-Escuredo, D., Torres Fernandez, Y., Bauer, J.M., Wadhwa, A., Castro-Correa, C., Romanoff, L., Lee, J.G., Rutherford, A., Frias-Martinez, V., **Oliver, N.**, Frias-Martinez, E. and Luengo-Oroz, M.

[12]Younes Mounmy, Vanessa Frias-Martinez and Enrique Frias-Martinez, "Characterizing Social Response to Urban Earthquakes using Cell-Phone Network Data: The 2012 Oaxaca Earthquake", Third Workshop on Pervasive Urban Applications @Pervasive'13, Zurich,Switzerland 2013.

[13]Enrique Frias-Martinez, Graham Williamson and Vanessa Frias-Martinez, "Simulation of Epidemic Spread Using Cell Phone Call Data: H1N1 Case Study", NetMob 2013 (Poster).

[14]Vanessa Frias-Martinez, Alberto Rubio and Enrique Frias-Martinez, "Measuring the Impact of Epidemic Alerts on Human Mobility using Cell-Phone Network Data", Second Workshop on Pervasive Urban Applications @Pervasive'12, Newcastle, UK, 2012.

[15] Vanessa Frias-Martinez, Cristina Soguero, Malvina Josephidou and Enrique Frias-Martinez, "Forecasting Socioeconomic Trends With Cell Phone Records", 3rd ACM Symposium on Computing for Development, DEV'13, Bangalore, India, 2013.

[16] Enabling Humanitarian Use of Mobile Phone Data Montjoye, Y.A., Kendall, J. and Kerry, C. Issues in Technology Innovation Nov 2014
<http://www.brookings.edu/~media/research/files/papers/2014/11/12%20enabling%20humanitarian%20mobile%20phone%20data/brookingstechmobilephonedataweb>

[17] Waiting on Hold. Ebola and Big Data. The Economist Nov 2014
<http://www.economist.com/news/science-and-technology/21627557-mobile-phone-records-would-help-combat-ebola-epidemic-getting-look>