

World Bank

Jennifer Ferreira

Accepted manuscript PDF deposited in Coventry University's Repository

Original citation:

Ferreira J. (2017) World Bank. In: Schintler L., McNeely C. (eds) Encyclopedia of Big Data.

<http://doi.org/10.1007/978-3-319-32001-4>

Publisher: Springer

Copyright © and Moral Rights are retained by the author(s) and/ or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This item cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder(s). The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holders.

World Bank

The World Bank, part of the World Bank Group established in 1944 is the international financial institution responsible for promoting economic development and reducing poverty. The World Bank has two key objectives: to end extreme poverty by reducing the proportion of the world's population living on less than \$1.25 a day; and promoting shared prosperity by fostering income growth in the lowest 40% of the population.

A core activity for the World Bank is the provision of low interest loans, zero to low-interest grants to developing countries. This could be to support a wide range of activities from education and health care, to infrastructure, agriculture or natural resource management. In addition to the financial support, the World Bank provides policy advice, research, analysis and technical assistance to various countries in order to inform its own investments, but ultimately to work towards its key objectives. Part of its activities relate to the provision of tools to research and address development challenges, some of which are in the form of providing access to data; for example the Open Data website which includes a comprehensive range of downloadable data sets related to different issues. This shows its recognition of the demand for access to quantitative data to inform development strategies.

A significant amount of the data hosted and disseminated by the World Bank is drawn from national statistical organizations, and it recognizes that the quality of global data therefore is reliant on the capacity and effectiveness of these national statistical organizations. The World Bank has ten key principles with respect to its statistical activities (in line with the Fundamental Principles of Official Statistics and the Principles Governing International Statistical Activities of the United Nations Statistical Division: quality; innovation; professional integrity; partnership; country ownership; client focus; results; fiscal responsibility; openness and good management.

The world is now experiencing unprecedented capacity to generate, store, process, and interact with data; a phenomenon that has been recognized by the World Bank, like other international institutions. For the World Bank, data is seen as critical for the design, implementation and evaluation of efficient and effective development policy recommendations. In 2014, Jim Yong Kim, the President of the World Bank discussed the importance of efforts to invest in infrastructure, including data systems. Big data is recognized as a new advancement which has the potential to enhance efforts to address development, although it recognizes there are a series of challenges associated with this. In

2013 the World Bank hosted an event where over 150 experts, data scientists, civil society groups and development practitioners met to analyse various forms of big data and consider how it could be used to tackle development issues. The event was a public acknowledgement of how the World Bank viewed the importance of expanding the awareness of how big data can help combine various data sets to generate knowledge which can in turn foster development solutions.

A report produced in conjunction with the World Bank, *Big Data in Action for Development* highlights some of the potential ways in which big data can be used to work towards development objectives, and some of the challenges associated with doing so. The report sets out a conceptual framework for using big data in the development sector highlighting the potential transformative capacity of big data, particularly in relation to raising awareness, developing understanding and contributing to forecasting.

Using big data to develop and enhance awareness of different issues has been widely acknowledged. Examples of this include: using demographic data in Afghanistan to detect impacts of small scale violence outbreaks; using social media content to indicate unemployment rises or crisis related stress; or using tweets to recognise where cholera outbreaks were appearing at a much faster rate than was recognised in official statistics. This ability to gain awareness of situations, experiences and sentiments is seen to have the potential to reduce reaction times, and improve processes which deal with such situations.

Big data can also be used to develop understanding of societal behaviours. Examples include: investigation of twitter data to explore the relationship between food and fuel price tweets and changes in official price indexes in Indonesia; after the 2010 earthquake in Haiti, mobile photo data was used to track population displacement after the event; satellite rainfall data used in combination with qualitative data sources to understand how rainfall affects migration.

Big data is also seen to have potential for contributing to modelling and forecasting. Examples include: the use of GPS equipped vehicles in Stockholm providing real time traffic assessments which used in conjunction with other data sets such as weather can then be used to make traffic predictions; using mobile phone data to predict mobility patterns.

The World Bank piloted some activities in Central America to explore the potential of big data to impact on development agendas. This region has historically experienced low

frequencies of data collection for traditional data forms, such as household surveys and so other forms of data collection were viewed as particularly important. One of these pilot studies used google trends data to explore the potential for the ability to forecast price changes to commodities. Another study, in conjunction with the UN Global Pulse explored the use of social media content to analyze public perceptions of policy reforms, in particular a gas subsidy reform in El Salvador, highlighting the potential for this form of data to complement other studies on public perception.

The report from the World Bank, *Big Data in Action for Development* presents a matrix of different ways in which big data could be used in transformational ways toward the development agenda: using mobile data (e.g. reduced mobile phone top ups as an indicator of financial stress); financial data (e.g. increased understanding of customer preferences); satellite data (e.g. to crowd source information on damage after an earthquake); internet data (e.g. to collect daily prices); and social media data (e.g. to track parents perception of vaccination). The example of examining the relationship between food and fuel prices and corresponding change in official price index measures by using twitter data (by the UN Global Pulse Lab) is outlined in detail explaining how it was used to provide an indication of social/economic conditions in Indonesia. This was done by extracting tweets mentioning food and fuel prices between 2011-2013 (around 100,000 relevant tweets after filtering for location and language) and analyzing these with corresponding changes from official data sets. The analysis indicated a clear relationship between official food inflation statistics and the number of tweets about food price increases. This study was cited as an example of how big data could be used to analyse public sentiment, in addition to objective economic conditions. The examples mentioned here are just some of the activities undertaken by the World Bank to embrace the world of big data.

As with many other international institutions which recognise the potential uses for big data, the World Bank also recognizes there are a range of challenges associated with the generation, analysis and use of big data.

One of the most basic challenges for many organisations (and individuals) is gaining access to data, from both government institutions and the private sector. A new ecosystem needs to be developed where data is made openly available and sharing incentives are in place. It is acknowledged by the World Bank that international agencies will need to address this challenge by not only by promoting the availability of data, but promoting collaboration and

mechanisms for sharing data. In particular, a shift in business models will be required in order to ensure the private sector is willing to share data, and governments will need to design policy mechanisms to ensure the value of big data is captured, and is shared across departments. Related to this, there need to be considerations of how to engage the public with this data.

Thinking particularly about the development agenda at the heart of the World Bank there is a paradox: countries where poverty is high or where development agendas require the most attention are often countries where data infrastructures or technological systems are insufficient. Because the generation of big data relies largely on technological capabilities, relying on those who use or interact with digital sources may be systematically unrepresentative of the larger population that forms the focus of the research.

The ways in which data are recorded have implications for the results which are interpreted. Where data is passively recorded there is less potential for bias in the results generated, and likewise where data is actively recorded there is greater potential for the results to be more susceptible to selection bias. Furthermore, how data is processed into a more structured form from the often very large and unstructured data sets requires expertise to both clean the data and where necessary aggregate it (for example if one set of data collected every hour, and another every day). Then the media through which data is collected is also an important factor to consider. Mobile phones for example producing highly sensitive data, satellite images produce highly unstructured data, and social media platforms produce a lot of unstructured text which requires filtering and codifying which in itself requires specific analytic capabilities.

Then in order to make effective use of big data, those using it need to consider elements about the data itself. The generation of big data has been driven by advances in technology, yet these advances are not alone sufficient to be able to understand the results which can be gleaned from big data. Transforming vast data sets into meaningful results requires effective human capabilities. Depending on how the data is generated, and by whom, there is scope for bias, and therefore misleading conclusions. Then with large amounts of data there is a tendency for patterns to be observed where there may be none; because of its nature big data can give rise to significant statistical correlations. It is important to remember that correlation does not imply causation. Then just because there is large amount of data available, this does not necessarily mean this is the right data for the question or issue being investigated.

The World Bank acknowledges that for big data to be made effective for development there will need to be collaboration between practitioners, social scientists and data scientists in order to ensure the understanding of the real world conditions, data generation mechanisms, and methods of interpretation are effectively combined. Beyond this there will need to be cooperation between public and private sector bodies in order to foster greater data sharing and incentivize the use of big data across different sectors. Even when data has been accessed, in nearly all occasions it needs to be filtered, and made suitable for analysis. Filters, require human input, and need to be applied carefully as their use may preclude information and affect the results. Data needs to be cleaned. Mobile data is received in unstructured form of millions of files, which requiring time-intensive processing to obtain data suitable for analysis. Likewise, analysis of text from social media requires a decision making process to filter out suitable search terms.

Finally, there are a series of concerns about how privacy is ensured with big data, given that often there are elements of big data which can be sensitive in nature (either to the individual, or commercially). This is made more complicated as each country will have different regulations about data privacy which poses particular challenges for institutions working across national boundaries, like the World Bank.

For the World Bank, the use of big data is seen to have potential for improving, and changing the international development sector. Underpinning the ideas of the World Bank's approach to big data is the recognition that while the technological capacities for generation, storage and processing of data continue to develop this also needs to be accompanied by institutional capabilities to enable big data analysis to contribute to effective actions that can contribute to development, whether this is through strengthening of warning systems, raising awareness or developing understanding of social systems or behaviors.

The World Bank have begun to consider an underlying conceptual framework around the use of big data, in particular considering the challenges it presents in terms of using big data for development. In the report *Big Data in Action for Development* it is acknowledged that there is great potential for big data to provide a valuable input for designing effective development policy recommendation, but also that big data is no panacea. The World Bank has made clear efforts to engage with the use of big data, and has begun to explore areas of clear potential for big data use. However, questions remain about how it can support countries to take

ownership and create, manage and maintain their own data, contributing to their own development agendas in effective ways.

Jennifer Ferreira

Coventry University

See Also: Bank of America; Citigroup Inc; International Development; United Nations; United Nations Global Pulse; World Health Organization

Further Readings

Coppola, Andrea. Calvo-Gonzalez, Oscar., Sabet, Elizabeth., Arjomand, Natalia. Siegel, Ryan. Freeman, Carrie. and Massarat, Neisan. (2014) *Big Data in Action for Development*. World Bank and Second Muse. Available at:

http://live.worldbank.org/sites/default/files/Big%20Data%20for%20Development%20Report_final%20version.pdf

LaValle, Steve, Lesser, Eric, Shockley, Rebecca, Hopkins, Michael, Kruschwitz, Nina (2011). Big data, analytics and the path from insights to value. MIT Sloan Management Review 52(2): 21-31.

Lehdonvirta, Vili, and Ernkivist, Mirko. *Converting the virtual economy into development potential: knowledge map of the virtual economy*. InfoDev/World Bank White Paper, 1:5-17. (2011).

McAfee, Andrew. and Brynjolfsson, Erik. *Big data: the management revolution*. Harvard Business review, 90(10): 60-66. (2012).

United Nations Global Pulse. *Big Data for Development: Challenges & Opportunities*. UN, New York: NY. (2012)