



# Q&A **BIG DATA: A ROAD TO DEVELOPMENT**

 by Cristina Serra

*Big data has enormous potential to benefit human communities. But two African data experts say that scientists, policymakers and the public must work together to achieve that potential.*

**W**hat do traffic flow, epidemics, mobile phones and marketing campaigns have in common? They all generate huge amount of data. But to distill knowledge from the oceans of big data, it will be essential to develop skills in data collection, management and analysis.

By building that capacity, developing countries could better address local challenges, say Leonida Mutuku, CEO at Kenyan-based company Intelipro. And to encourage policy that supports capacity-building, scientists must seek partnerships with policymakers and the public, adds TWAS Fellow Tshilidzi Marwala, deputy vice chancellor at the University of Johannesburg.

“It is difficult to solve problems that you can’t quantify and the use of big data enables developing countries to determine the extent of these problems.”

Leonida Mutuku

Big data flow every day from all human activities, whether financial transactions, social media messages or climate sensors. Assembling, storing, analysing and sharing that information would make it easier for Least Developed Countries to address their own problems and to engage in regional and global research and policy decisions.

With so many other needs, can developing nations afford to make

the investments in education and technology? Mutuku and Marwala, in interviews with TWAS staff writer Cristina Serra, posed a different question: Can developing nations afford to miss this opportunity?

**Least Developed Countries have urgent problems to address – food and water, energy, sanitation, and primary education. They often have great difficulty responding to natural disasters. Why should they invest in big data?**

**Mutuku:** Big data is an enabling resource to solve challenges facing LDCs. It is difficult to solve problems you can’t quantify and the use of big





“The only way to build capacity in the area of big data is to educate the population.”

*Tshilidzi Marwala*

project that facilitates movement of experts to LDCs. I think LDCs should use this opportunity. With the recent development of massive online open courses, one is now able to learn about big data without travelling to the developed world.

**For scientific and economic growth of developing countries, how essential is it that policymakers understand the importance of investment in big data?**

**Mutuku:** National Statistics Offices [NSOs] should be at the core of the big data conversation in the context of development. NSOs, as the official sources of government data for resource planning, require increased investments and capacity to develop alternative and faster methods of collecting and analyzing demographic and economic data. Arguments for governments to embrace big data and open data have been towards the reduction and identification of corruption. This, however, is not a strong motivation for political will. Moving towards evidence-based governance is more likely to win political buy-in at the government level and in public institutions. Initiatives such as the African Data Consensus provide a roadmap for policymakers in investing in big data.

**Marwala:** It is essential that policymakers start to believe in big data. Policy should be evidence-based. Today the evidence is hidden in huge

data enables such countries determine the extent of these problems before setting a roadmap to solve them. For instance, the availability of gender data is very critical to ensure that women and youth are not left behind in critical interventions.

**Marwala:** Big data, the utilization of vast data located in the Internet for decision support, can be on weather patterns which have impact on food security, energy demand or water demand. So it will be unwise for any country to ignore it, especially intelligent analytics that can assist us to plan better in fields like energy and agriculture or to anticipate natural disasters. “Ignore big data at your peril!”

**Developing countries often lack the resources to compete in big projects. What kind of technology should they have? How can they build capacity with limited resources and access to technology?**

**Mutuku:** A core requirement for big data analytics is access to affordable

and fast Internet. Governments of developing countries are rolling out national fiber-optic backbones within their countries to increase access to reasonably priced broadband connectivity. This opens up opportunities for local cloud and related infrastructure that can manage big data projects with limited latency. Another investment developing countries should be making is in human capital: individuals and innovators who can work closely with governments to improve data collection methods and analysis capacity of public institutions.

**Marwala:** The starting point in big data regards the knowledge of mathematics, especially statistics, then computer programming and at the high-end artificial intelligence. The only way to build capacity in the area of big data is to educate the population. This can entail importing skills to come and teach in areas such as statistics, artificial intelligence and computer science, which are essential for big data. TWAS has a



## PROFILES

**Leonida Mutuku** is a technology researcher, data scientist, entrepreneur, and investor. She is the co-founder and CEO at Intelipro, Kenya, a consultancy that builds data products and analytical tools for financial and retail organizations. She has done research on open data and big data in developing countries. She led the Research and Data Science teams at iHub Research in Keyna (2011-2015), guiding a team in groundbreaking research on the role of open data in increasing access to information, and the potential for technologies and big data to improve governance and urban resilience. Mutuku is one of the three panelists who will discuss big data at the TWAS roundtable during Trieste Next 2016, an international science festival that runs 23-27 September in Trieste, Italy.

**Tshilidzi Marwala**, elected to TWAS in 2010, is currently the deputy vice chancellor in research, innovation, postgraduate studies and the library at the University of Johannesburg. He was the executive dean of the faculty of engineering and the built environment at the University of Johannesburg. He is a member of the ICT Advisory Council of the South African government. Marwala was the TWAS representative in Science International working group that prepared the "Open Data in a Big Data World" accord urging global endorsement of principles and policies for open access to big data.

databases in the Internet. Without big data, such knowledge will remain hidden from policymakers and this can only limit their effectiveness. For example, power stations can only work better when machines help human beings in making better decisions, and these machines use big data. Big data enables countries to anticipate interstate conflict and help them make decisions in the presence of missing information.

### What is the interest of developed countries to help developing countries acquire the capacity to use big data?

**Mutuku:** Developed countries are large aid donors to developing countries. Big data use in poor countries helps them move towards self-sustainability and less reliance on donor funding through proper planning and use of resources at their disposal. Secondly, with improved economic conditions for developing countries as a result of big data use, new markets or opportunities

for trade partnerships open up. The concern is that developed countries are pushing developing countries to collect better data for self-serving purposes.

**Marwala:** New markets are essential for developed countries. Unlocking the economic activities of developing countries makes business sense for developed countries.

### Is there a role for science diplomacy in helping to build global big data capacity?

**Mutuku:** Science diplomacy offers very useful collaborations towards not only building big data capacity but also achieving breakthroughs in solving global challenges. It should be used as an augmentation to help developing countries achieve their own priorities and not as just another opportunity to receive donor funding.

**Marwala:** Absolutely. Until an average person on the street knows about big data and its economic, political

and social implications, then a great potential in countries will remain unrealized. Science diplomacy means we should educate the average person to understand science and technology and, in this particular case, big data.

### Choose one of the 17 UN Sustainable Development Goals and explain why big data will help developing countries to achieve the goal by 2030.

**Mutuku:** Take for instance goal No. 2: No hunger. There is no reason why countries cannot feed their own populations, as hunger is a result of mismanagement of resources and poor planning. Big data offer opportunities that can be used to drive the achievement of the SDGs, like:

1. Commodity exchanges and pricing: Local farmers can use big data analytics to lock down great pricing of their crop at an early stage.
2. Precision agriculture: Use of big data and predictive analytics can enable farms to accurately manage and optimize resources, ensuring no wastage and high yields at optimal costs.
3. Insurance: Big data enables accurate prediction and planning for droughts and floods to reduce famine. Similarly, insurance companies are able to create packages to compensate poor farmers for reduced yields due to weather fluctuations.

**Marwala:** Let's take goal No. 3: good health and well being. Much of the information that people require to lead a good life lies buried in the Internet. Today, especially in Africa, there has been a growth in mobile telephone use. If these devices can be used to disseminate information on health, then life expectancies can be increased. If it can be used to give doctors in rural areas access to relevant information, then delivery of medical services will be increased. ■