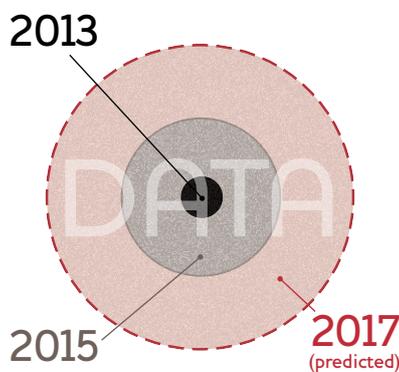
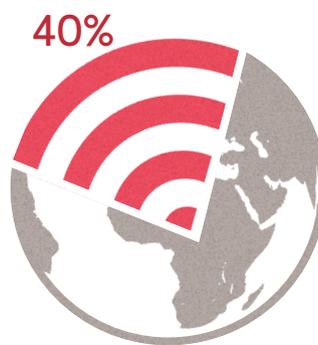


Ensuring Developing Countries Benefit from Big Data

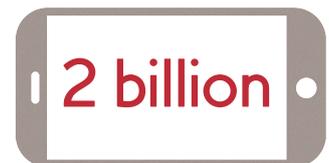
Big data is already creating a big impact. Some herald it as the new ‘data revolution’, others worry it is a tool for manipulation, while many in development cite it as key to successfully implementing the Sustainable Development Goals. Major decisions on big data are being taken now, but much more needs to be done to ensure the risks are minimised and the benefits shared equitably and not just among rich individuals, corporations and developed nations.



90 per cent of data in existence were created in the last two years, and the quantity is doubling every two years



40 per cent of the global population use the internet every day



By 2016, two billion people will have an internet-enabled smartphone

Influencing key ‘big data’ decisions

We define big data as: *our growing ability to generate, manage, analyse and synthesise data to create and destroy different forms of value.* For its evangelists, big data will be the fuel that drives the next industrial revolution, radically reshaping economies and societies, and reaching into every aspect of life. Others worry that it will be used to create a ‘surveillance state’, or that mega-corporations will be able to ‘see’ everything we do and use this information to manipulate us. None of this is certain. What we do, and don’t do, over the next few years will be crucial in shaping the data future we get.

Key decisions are being taken now. The European Court of Justice recently ruled that the ‘safe harbour’ agreement, which allowed firms such as Google to hold data on their European customers in the US, was invalid. This has opened up fundamental questions of how personal data are used. Issues of data privacy and use are also being discussed as part of trade negotiations between the US and EU (Transatlantic Trade and Investment Partnership [TTIP]) and between 24 high- and middle-income countries around the world (Trade in Services Agreement [TISA]). The outcome of these negotiations will shape big data impacts

“We define big data as: *our growing ability to generate, manage, analyse and synthesise data to create and destroy different forms of value.*”

for years to come. While still unclear, the impact of the European ruling on negotiations on these issues is likely to be significant.

The fact that agreements like TTIP and TISA are being negotiated in secret makes it impossible for citizens to engage. That the long-term implications of big data are generally not well understood makes this harder still. Increasing understanding of these issues, and shining a light into the opaque world of decision-making in this area, is an urgent task, particularly when many are arguing that the 'data revolution' will be an essential component of implementing the Sustainable Development Goals.

The nature of big data impacts

One way of thinking about the potential impacts of big data is in terms of different types of 'value' that could be created, or in some cases destroyed. We identify four broad areas:

- i. **Economic:** The monetary value created for companies through new or increased profits, for people (income) or governments (tax revenue). Economic value is destroyed when companies fail, livelihoods are lost or tax revenues are reduced.
- ii. **Human development:** The value created through advancing health, education, or other development outcomes, and the individual or social value that may be created or lost through new forms of communication and interaction.
- iii. **Rights:** The value to individuals or societies of freedom of speech, expression, and movement, and of rights to privacy and protection under the law. These types of value can be destroyed by invasions of privacy and the curtailing of other freedoms.
- iv. **Environment:** The environment has intrinsic value and creates value through the utility people gain from its resources. Reducing strain on resources creates environmental value, and vice versa.

Understanding these potential effects in developing countries, and identifying how policy could be used today to influence these positively is critical.

The value of big data in developing countries

Economic value

Big data is directly creating entirely new industries, infrastructure and markets. This may have particular importance in developing countries, whose firms can compete directly for data analysis work across the world. Data-intensive companies from developed countries can now also outsource data analysis to countries where labour costs are low but skills are high. Kenya and Rwanda are already known for their IT capabilities, for example. Platforms such as Amazon's 'Mechanical Turk' outsource small, low-skilled jobs

involving simple data manipulation tasks, which only require a basic internet connection. Less positively, the geographical concentration of data scientists in some countries, particularly the US, may exacerbate the 'brain drain' from developing countries.

For indirect economic effects, real-time data monitoring and information sharing will create efficiency gains in many aspects of manufacturing, transport and logistics. As more devices and simple products are equipped with smart sensors, increasingly efficient automation processes will be enabled. Big data-informed organisational restructuring should also result in significant efficiency gains. While this increased efficiency will create value for firms by boosting profits, it may destroy value for some workers as jobs are lost.

Perhaps the most visible applications of big data so far have been in marketing, with advertising increasingly tailored to individuals based on their online activities. For some, this is a value-creating process, saving them time by editing the 'choices' presented to them about potential purchases. For others, however, it is intrusive and even unsettling to be so closely monitored.

Although developing countries could, in principle, compete directly in the big data economic space, this seems highly unlikely to succeed given the 'digital divide' that exists in skills and infrastructure. This suggests that the overwhelming bulk of profits and returns created will go to today's technologically advanced countries, specifically to the owners of big data firms, and their employees.

While the net jobs impact is uncertain, the creation of large numbers of new jobs in big data-related industries seems unlikely. There may also be less physical outsourcing through greater use of data-driven automation in manufacturing. At the same time, the internet has made it possible to outsource data analysis tasks globally. Concerns are that many of these tasks will be low-skill, competition to secure them will be high, and they will only be viable as long as they are cheaper than automating these tasks – the costs of which are falling all the time.

Human development

The increased ability of people to connect and interact is an important direct impact of big data. Whilst the US remains the biggest market in most areas, the fastest growth rates are being seen in the developing world. Indonesia, for example, has the most Twitter users globally.

Some argue that big data will mean the end of the scientific method. Traditional modelling, for example, makes use of a small sample of observations to make predictions. Big data allows much larger sets of data, potentially increasing reliability significantly.

Critics argue that bias is inevitable. Analysing Twitter feeds, for example, biases results to Twitter users, who are not evenly spread across socioeconomic and racial groups.

A defining feature of big data is the ability to synthesise different data sets in new ways, such as the use of remotely sensed and crowdsourced data to 'map' problems of many types, such as the spread of malaria, influenza, or dengue fever. Remote and real-time patient monitoring is already being used in development interventions for health and nutrition.

Big data can potentially help teachers and administrators through data-driven decision-making and mining of large amounts of data. Education can also increasingly be delivered remotely cheaply or for free. Students in developing countries can now increasingly access massive open online courses, for example.

A more fundamental issue for all countries is that big data may undermine support for public health and education services. It is claimed that big data techniques will increasingly be able to predict individual outcomes accurately but the basis for pooling risk through public health services, for example, is that individuals do not know when and how much they will need these services. If uncertainty in these areas is reduced or eliminated, those with the resources to pay privately for only what they need will have stronger incentives to do so, undermining public support for collective 'insurance' in areas like health.

Potential benefits in developing countries require reliable internet connections and online devices. They also require subsidised public provision or affordable business models capable of delivering quality services at low prices. More specifically, for monitoring systems to work effectively in developing countries, the current bias towards English language data needs to be addressed. Algorithms that trawl social media 'exhausts' for information largely focus on the English speaking world. Finally, whether the remote delivery of services undermines domestic capacity will be influenced by whether these services are integrated with local providers or not.

While the potential human development benefits of big data are undoubtedly greater than economic benefits, realising this potential is far from guaranteed.

Rights

New hardware, software and analytical techniques have vastly increased the potential for surveillance by

government and corporations. These same innovations have opened up new ways for civil society groups to organise to advance or defend their rights, and also to hold governments to account by increasing transparency over their activities. Understanding what will shape outcomes in this area, and what policy can do to influence this, is an urgent task.

The extent to which corporations and governments are able to invade privacy depends on the restrictions placed on their activities. Given the increasing use of automated decision-making based on black-box algorithms, this is not just a matter of human activities. Ensuring that these algorithms do not have unconscious biases towards particular groups, as some have been found to have, requires careful design. Given how new these techniques are, this is only likely to emerge after much trial and error, inputs from a variety of sources, transparency and mutual learning. Unfortunately the current secrecy surrounding how algorithms are designed, which is often seen as commercially sensitive, makes this impossible.

Whether the worst fears about privacy actually happen will be strongly influenced by whether people have rights to see what data are held about them, and to control how they are used. For big data to empower civil society groups rather than a surveillance state, therefore, governments need to be restricted in their ability to monitor citizens, but citizens empowered to monitor governments. Some take a different view, raising concerns that preventing government from monitoring citizens creates security risks, as terrorist groups – for example – can and do organise through the same techniques.

Developing countries face particular challenges with respect to rights. While the need to hold governments to account is often high, historical protection for civil liberties is not encouraging in many cases, and concerns over privacy may have less prominence than they could have. The worst of all worlds is one where government can 'see' their citizens with the mosaic effect, but information on government activities remains closed, and many citizens have only limited, 'tethered' access to information, edited by their corporate 'provider'. It is interesting to note the increasing number of initiatives that offer internet access to people in developing countries in exchange for targeted advertising and the right to use their data. As one commentator put it, the risk is that 'the rich pay for their privacy while the poor pay with their privacy.'

“While the potential human development benefits of big data are undoubtedly greater than economic benefits, realising this potential is far from guaranteed.”

Policy recommendations

Global-level policies:

- Establish a United Nations (UN) panel of social scientists, ethicists, legal and technical experts, to design a Declaration of Data Rights, to balance privacy with the potential benefits data could bring for people from all countries. It should enshrine citizens' rights to access data on their government's activities in the process and a citizen's right to see and control the information held about them, by governments and corporations.
- Research whether data-related firms need to be very large to generate the benefits associated with big data, or if a greater number of smaller, 'networked' firms could produce the same outcomes. Use results to inform whether large incumbent firms should be broken up, or regulated more strictly to ensure the public benefits of advances in data are maximised.
- Increase funding for public research into the implications of the increasing use of automated decision-making and learning algorithms, and regulate to require corporations to increase transparency of their activities in this area.
- 'Bespoke' health and education services developed from real-time monitoring and online services are unlikely to be widely affordable based on commercial models. This suggests the need for public sector provision of health and education services; and donor support for this, as well as concessional financing and social enterprise models, will need to continue as 'bespoke' health and education services developed from real-time monitoring and online services are unlikely to be widely affordable based on commercial models.
- Undertake research on the feasibility of creating a 'marketplace' where people could sell as much or little of their personal data as they choose. Particular attention would need to be paid to whether this can be done in a non-exploitative way, given current inequalities of income and wealth.
- Practice and encourage a presumption in favour of freedom of speech and association for all groups, within clear legal boundaries.
- Require companies from developed countries to employ the same approach to data privacy in all countries that they operate in as they do at home.

Developing country-specific policy:

- Encourage all countries to sign a UN Declaration of Data Rights, and develop robust policies to protect citizens' rights without preventing the potential benefits of big data from being realised.
- Increase the supply of donor funds, international private investment and domestic public and private resources devoted to ICT infrastructure, maximising the access citizens have to internet-enabled devices, the internet and affordable power.
- Increase investment in statistical and computing skills to reduce the 'digital divide' and realise potential benefits of big data.
- Generate revenues to invest in 'smart' urban design and transport infrastructure, and 'smart' health and education systems. In some cases this may be commercial finance (e.g. some forms of infrastructure), in others donor finance and domestic public resources will be needed, either to leverage private finance, or to provide concessional or grant financing directly.
- Where feasible, develop 'digital industrial policy' to create competitive domestic firms in big data 'niches' and prevent 'brain drain'. In countries where it is simply not realistic to create globally competitive firms in this space, resources should not be wasted in attempts to do so. Rather, the focus should be on ensuring that domestic regulation is well designed and implemented.
- Address the English language bias of big data early warning programmes, and increase their ability to tap into grass-roots real-time data sources.

Further reading

Independent Expert Advisory Group on a Data Revolution for Sustainable Development (2014) *A World That Counts: Mobilising the Data Revolution for Sustainable Development*, www.undatarevolution.org/wp-content/uploads/2014/12/A-World-That-Counts2.pdf

Spratt, S. (2015) *Big Data and International Development: Impacts, Scenarios and Policy Options*, IDS Evidence Report 163, Brighton: IDS

Mayer-Schönberger, V. and Cukier, K. (2013) *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, New York: Eamon Dolan/Houghton Mifflin Harcourt

Privacy International (n.d.) *What is Big Data?*, www.privacyinternational.org/node/8

Credits

This *IDS Policy Briefing* was written by **Stephen Spratt**, Research Fellow and edited by **Hannah Corbett** and **Sophie Robinson**. It was produced as part of the IDS Policy Anticipation, Response and Evaluation programme, supported with UK aid from the UK Department for International Development.

The opinions expressed are those of the author and do not necessarily reflect the views of IDS or the UK government's official policies.

Readers are encouraged to quote and reproduce material from issues of *IDS Policy Briefings* in their own publication. IDS requests due acknowledgement and quotes to be referenced as above.

AG Level 2 Output ID: 254

© Institute of Development Studies, 2015
ISSN 1479-974X



IDS Policy Briefings are published by the Institute of Development Studies and aim to provide high quality analysis and practical recommendations for policymakers on important development issues.

To subscribe: www.ids.ac.uk/idspolicybriefings

Institute of Development Studies, Brighton BN1 9RE UK
T +44 (0) 1273 606261 F + 44 (0) 1273 621202 E ids@ids.ac.uk W www.ids.ac.uk
twitter.com/IDS_UK #idspolicy facebook.com/idsuk