

Free Access to Primary Data Should Be a Right

by Rafael Guerreiro Osorio, International Poverty Centre

Few things are as important for development as the availability of data and unfettered access. Without data there can be no needs assessment. It will scarcely be possible to design effective policies or to implement, monitor and evaluate them. Data are not important only for governments. Democracy implies transparency of government actions and the accountability of governments to society. The availability of data from reliable sources is a step towards these goals.

Countries differ considerably in data production, as well as in their data dissemination policies. Both data production and dissemination influence countries' capacities to undertake development planning. The availability of data and free access also determine the extent to which the electorate can evaluate the performance of governments and their policies. Lack of capacity to produce data might have many excusable causes, but restrictive policies of data dissemination have none. As long as the anonymity of individuals and firms is guaranteed, all data should be available in the public domain.

Governments are usually the main producers of primary data. The most common primary data sources are administrative records, censuses and surveys. It is from such primary data that indicators are calculated. Many governments disseminate indicators, but not the primary data. Governments that withhold access to primary data impose restrictions in many ways. Some simply do not grant access at all; some grant access only to government officials; some grant access only when commissioning a consultancy or receiving something in exchange; some charge very high fees. By restricting access to primary data, governments ensure that their indicators will not be contested. Moreover, policy shortfalls that they do not want exposed are kept secret.

For a long time, there were technological limitations on the dissemination and use of data. Data dissemination was costly, as were analytical tools. In some cases, lack of technology and high costs may have been used as pretexts for restricting the availability of primary data. Technological progress has changed this, and today there are almost no limits on data dissemination. In most countries, cheap personal computers and internet access are becoming widespread, endowing individuals with processing power that once was associated with supercomputers.

Governments lose by restricting access to primary data. Worldwide, an increasing number of researchers are eager to acquire such data.

Be it because academic competition requires them to publish papers, or because they want a say in decision-making, access to primary data allows them to test their hypotheses about development policy options. Each analysis becomes a potential free consultancy with particular views. So instead of commissioning a costly consultancy and receiving a single analysis, countries that disseminate primary data can acquire a full menu of policy options for free.

Fortunately, some governments have recognized the advantages of disseminating primary data and have begun to allow access to datasets. But it is not enough to make the data available if the information is not usable. Primary data dissemination packages must include data itself and metadata—the information on how the data were produced. The data must be disseminated in standard, non-proprietary file formats. Metadata should include descriptions of the data files, copies of the forms or questionnaires used in data collection, copies of the form-filling instructions or interview manuals and, in the case of surveys and censuses, comprehensive methodological documents.

There are many good examples of primary data dissemination. The outstanding one is that of the Demographic and Health Surveys (www.measuredhs.com). Statistics South Africa has done a remarkable job in making many of its surveys freely downloadable, with fair metadata (www.statssa.gov.za). The information technology department of Brazil's health ministry is a success story in the dissemination of primary data from administrative records (www.datasus.gov.br).

In Latin America, many statistical bureaus have started to offer free dataset downloads from household surveys: those of Bolivia, Ecuador, Honduras, Mexico, Peru and Uruguay, for instance, already do this. Some still have to work on the comprehensiveness of the metadata, particularly for older data, but these countries' initiative is commendable.

The more people have access to primary data, the more analyses it will generate; more ideas will emerge, and the quality of the debate on development options will improve. Democracy will benefit from the greater accountability of government actions and the increase in plural participation. Furthermore, primary data should not be seen as the property of governments, but as a societal asset. Free access to primary data should be seen as a fundamental right, not as a concession.