

Big Data, Big Impact



June Gomes
One East Midlands

What is Big Data?

- “one byte more than you are comfortable with”
- Big data – using diverse sources of data sets, some of which might be very large and employing mathematical techniques to extract new information from them.
- Open data - Datasets that are made accessible in non-proprietary formats under licences that permit unrestricted re-use. Open government data involves governments providing many of their datasets online in this way
- IT company IBM estimates 2.5 quintillion bytes of data are created every day

Do we need to crunch the numbers?

- “It is easy to keep on doing things you do because that's what you've always done. Data allows you to see if it's the right thing to do”, (Duncan Ross, Datakind UK)
- Experience + data =
 - shows what you are doing works
 - Shows you how to do things more effectively
 - Shows you how to do more with the resources you have
- Private sector been analysing data for years. E.g. Tesco Clubcard initiative collects data on its customers spending habits and uses it to inform the business.

Benefits of using open data

- Existing open government data can be used by charities/VCS:
 - to add value to their work,
 - to target services better,
 - to improve advocacy and fundraising
 - to support knowledge sharing and collaboration between different charities and agencies
 - to demonstrate impact

Case Study Macmillan Cancer Support



- Macmillan has been able to analyse NHS data sets, including cancer registration and mortality to create a desktop tool for commissioners to plan cancer care in advance in any area of the UK.
- The model developed by the charity predicts how many people will be diagnosed with cancer, what types they will have, how many will die and what their needs will be in 1-3 years.

Barnsley Hospice



Barnsley Hospice provides an example of how data can be used to improve fundraising. By linking the postcodes from their donor database to local council information on the socio-economic profile of neighbourhoods, they are able to tailor their fundraising strategy for each area; for example, by marketing small pledges for lower income areas and corporate giving where businesses are located

Barriers

- Realistic about the sectors knowledge, skills and capacity to work with data
- Understanding how and where to access big data sets.
- Measuring impact can be hard and a bit frightening for some VCS organisations.
- Do all VCS organisations really want to know the truth about their impact, and do they want others to know
- Will anyone listen to the results anyway or is it basically a pointless exercise?
- Do the benefits outweigh the costs?

Types of big and open data sets

- <http://data.gov.uk/> - There are over 9,000 datasets available, from all central government departments and a number of other public sector bodies and local authorities.
- Councils have been asked to publish info on spending of more than £500 in an open format online.
- The Justice Data Lab has already been launched by the Ministry of Justice, working with NPC to explore how charities can use data on reoffending to assess the outcomes of their programmes. <http://www.justice.gov.uk/justice-data-lab>

Useful places to look for data

- <http://www.hscic.gov.uk/searchcatalogue> - UK health and social care data
- <http://data.ncvo.org.uk/category/almanac/> - NCVO Almanac a directory of useful sources of data on the voluntary sector.
- The Global Value Exchange - is an open source data resource that relies on contributions from real practitioners. This “bottom up” approach allows the real experiences of stakeholders and organisations to be heard by others. <http://www.globalvaluexchange.org/>
- Data Unity is an open source web tool which lets you explore and visualise data then share discoveries with others. The tool promotes Open Data, where you can query and publish data from within your organisation and outside. <http://dataunity.org/>

Useful places to look for data

- This unit cost database (http://data.gov.uk/sib_knowledge_box/toolkit) brings together more than 600 cost estimates in a single place, most of which are national costs derived from government reports and academic studies. The costs cover crime, education and skills, employment and economy, fire, health, housing and social services. The costs can be used by local commissioners, charitable organisations and social enterprises to inform:
 - proposals for new interventions or the redesign of existing public services and
 - Feasibility studies and evaluations

What if we don't crunch the numbers?



- Are we in danger of losing contracts/funding/business to private sector/university organisations who can crack the data and use it to present a strong business case?

Activity

- What kind of data would be useful to VCS infrastructure/ or VCS organisations?
- Do you think this data is available?
- What would you need to do to get the data?
- What would you use the data for?

Where to get help with understanding and analysing data?

- Datakind and the Open Knowledge Foundation run courses to help people gain data skills and are helping charities to get to grips with their data.
- Clinks and NPC's have a guide that introduces you to the increasing number of software tools that may help you to better collect and analyse data and raises some issues to consider.

http://www.clinks.org/support-evaluation-and-effectiveness/practicalities-research?utm_source=LINX+322+-+3+Apr&utm_campaign=LINX322&utm_medium=email