The Guardian seminar in association with TechSoup Global

Getting to grips with big data

Charities and non-governmental organisations may be missing a trick by not making effective use of their information. What data do they need and how should they use it? SA Mathieson reports on a recent seminar

hen the chancellor, George Osborne announced plans in . March's budget to cap tax relief

on charitable donations, he was met with widespread protests. The proposals were subsequently shelved, but only after months of lobbying from charitable organisations.

But could they have opposed Osborne'splans more effectively, earlier, if they had made better use of data? Karl Wilding, head of policy and research at the National Council for Voluntary Organisations (NCVO), believes so. According to Wilding, charities should have been able to provide clear analytical evidence of how they used tax breaks for beneficial purposes. "That's the sort of stuff that is the real fuel for much more effective lobbying and advocacy," he said, but added: "We hardly have any data."

Wilding was speaking at a seminar organised by the Guardian, in association with TechSoup Global, a not-forprofit organisation that helps charities improve their use of technology and social media, to explore how charities can harness the power of data to convince people of their cause, improve their management and empower those they support. Along with a panel of experts, the seminar was attended by an invited audience from a range of UK charities and not-for-profit organisations who added to the discussion.

As well as for advocacy, charities should also use data to save money, such as by publishing data on how they operate in real-time, with regulators taking that information to produce annual reports, Wilding said, adding that NCVO "probably spends £100,000 extracting data from PDFs of charities' accounts".

"We're thinking of all individual NGOs and civil society organisations as people that can contribute to the data stream," TechSoup Global's co-chief executive officer, Marnie Webb, told the audience.

She said that conversations about "big data" - a term used for very large amounts of unstructured information often get highly technical: "People start saying 'well-formed XML' and 'Hadoop', like those are words that are meant to mean something to normal human beings." She added that the important point is that tools now exist to process big data. "It's incredibly important for us, those that work in civil society, to get the information out of our heads, out of the heads of our constituency, unlocked from our computers and the PDFs and the nice annual reports we hand around in paper copies. We have to make it available so it's part of the decision-making structure of our communities."

Using big data



Attendees at the event discussed the various uses of data, including improving management, saving money and convincing people of a charity's cause **Photographs: Sam Freidrich**

Coplin, director of search for Microsoft's Bing search service. Microsoft is a long time supporter of TechSoup Global, which is funded mainly by IT companies and charitable foundations. Coplin said that he loves big data for three reasons: it provides insights unavailable without its use; it helps organisations, businesses and people to undergo exaptation - in other words, find new uses for what they already have, in the way that birds initially evolved feathers to keep them warm, then used them to fly; and it supports serendipitous discovery.

Coplin said that Microsoft had exploited big data techniques when designing the Kinect entertainment console, which is driven by users' gestures watched on a camera. Rather than decide what gesture should trigger which process, the company released trial devices "We tracked people's movement and usage," he said, then analysed vast amounts of resulting data to decide how the devices would interpret movement. He said that big data presents challenges around privacy of personal information, even when data has supposedly been anonymised. But he added that in general, when tools are made available for free, its use is empowering for both organisations and individuals: "The future rock stars of our world will be data analysts."

In response to a question, from an audience member who worked for a community voluntary service, on the costs and problems involved in processing data based on geography - something vital when dealing with councils - Coplin said that this was a government issue, not a technology one, and was deeply frustrating: "As a society, we could do some wonderful things if we agree that 80% of everything done in a local authority is probably the same." Councils should standardise whenever possible, he added.

On whether big data is more useful for advocacy or improving management, TechSoup Global's Webb said they should go together. She said that the city of Denver, Colorado, used data analysis that found that the best way to help homeless people was to give them an apartment - as long as they reported daily to a social worker - because this produces an overall saving through much lower emergency medical costs. "The data actually changed the way they managed these systems," she said, but added: "There was huge community resistance, so the advocacy and the new management practices had to go together." Data analysis that can convince an organisation to change can also be used to convince those outside the organisation to support that change. Van Haver added that politicians use sets of data to build policies and make arguments. This provides all the more reason for charities to do their own data collection and analysis: "If they don't hear the reality because we're not capturing it, then how can we really truly influence, and have that challenging conversation with an MP? A questioner from the audience who works on international aid said he knew of a project in the Democratic Republic of the Congo where staff spent half of their time putting information into a database. Microsoft's Coplin replied that organisations need to "try to get to a place where you spend more time thinking about what you're going to do with it [data], than getting it cleaned and in the right format and in the right place". Following group discussions at the seminar, participants reported that charities currently tend to use structured information in databases rather than big data, and quite often collect information and then don't make any use of it, or keep it within the organisation when it could be shared. They also raised the issue that useful data often

Resources

• Techsoupglobal.org is a nonprofit network that is working to develop and share innovative technological solutions to social challenges.

• Guidestarinternational.org is a charity that builds detailed reports on civil society organisations.

• **Ctt.org** provides qualifying charities with donated and heavily discounted software licences, communications and online payment services.

• Guardian.co.uk/data has collections of data, often open for further use.

The ideal, Webb said, is that charities and not-for-profit organisations collect lots of usable data as a by-product of their normal work, rather than it being a time-consuming extra job. She added that at present, the smartest use of big data is business-related, such as to target advertising: "But I'd hope that's not all we're doing with this."

Paul Van Haver, TechSoup Global's director of global data services, said that big data has three qualities: variety, volume and velocity. Relevant information is no longer held only in structured databases - it is in all kinds of documents, in logs that track behaviour and within social media services such as Twitter. Furthermore, the amount of data available is enormous, and it changes fast. "Luckily, the technology has progressed as well to help us actually do something meaningful with that data," he said.

Those attending heard how Ushahidi, an NGO that provides online services for sharing information, makes such meaningful usage as easy as possible. It was developed to track violence in Kenya followingthe2008election-thenamemeans "testimony" in Swahili - and users can contribute through text message, Twit-ter, email or web. "We build our platforms to integrate with the cheapest, simplest device out there," said its strategy director Nathaniel Manning. "The technology can end up being quite easy - it's how you implement." Ushahidi was also used during the Japanese earthquake last year, for monitoring elections in several countries and by the Guardian to collect reports of last summer's riots in England.

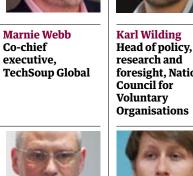
The event also heard from Dave

David Mills (Chair) Editor, Guardian voluntary sector network

On the panel



Paul Van Haver Director of global data services **TechSoup Global**



Dave Coplin Director of search, Bing

foresight, National

Nathaniel Manning Strategy director, Ushahidi

• Madwdata.org.uk advises on how to use open data from UK public bodies.

• Ushahidi.com curates eye-witness reports through text message, email, web or Twitter.

costs significant amounts of money, but that charities also try to sell information themselves to raise funds.

"What do we really want to collect?" asked Richard Craig, chief executive of the Charity Technology Trust, summarising his group's comments. "Maybe we're asking the wrong question, maybe we need to work out what questions we want answered in order to determine that." Charities could use other organisations' information, such as consumer and government data, and should aim for the serendipity discussed by Coplin, Craig suggested. He added: "Maybe we should be collecting more data than we need, because we don't know what we're going to stumble upon when we do the analysis work."



Seminar report commissioned by Seven Plus and controlled by the Guardian **Discussion** hosted to a brief agreed with TechSoup Global Paid for by TechSoup Global Contact Ben Hicks on 0203-353 3701 (ben.hicks@guardian.co.uk) For information on roundtables visit: guardian.co.uk/sponsored-content

