How social media data can improve people's lives—if used responsibly

April 13 2017, by Stefaan G. Verhulst

A woman participates in a community mapping exercise in Malawi's Chikwawa and Nsanje districts. Credit: emirhartato/flickr, CC BY-SA

In January 2015, heavy rains triggered unprecedented floods in Malawi. Over the next five weeks, the floods displaced more than 230,000 people and damaged over 64,000 hectares of land.

Almost half the country was labelled a "disaster zone" by Malawi's government. And as the humanitarian crisis unfolded, relief agencies,

such as [the Red Cross](#) were faced with the daunting task of allocating aid and resources to places that were virtually unrecorded by the country's mapping data, and thus rendered almost invisible.

Humanitarian workers [struggled to navigate in many of the most affected areas](#), and one result was that aid did not necessarily reach those most in need.

To prevent similar knowledge gaps in the future, researchers, volunteers and humanitarian workers in Malawi and elsewhere, have turned to an unlikely partner: Facebook.

In 2016, as part of its "Missing Maps" project, the Red Cross accessed Facebook's rich population density data to find and map people who were critically vulnerable to natural disasters and health emergencies, but remained unrecorded in existing maps.

During local [Mapping Parties](#), volunteers in Malawi used Facebook's satellite and population data, in addition to other satellite imagery, to trace roads, houses, and water points across Malawi's communities.

Two years later, Missing Maps in collaboration with Facebook has [identified more than 2,000,000 people in Malawi](#), allowing aid and relief organisations to better plan projects in Malawi's disaster prone areas.

Disasters kill [nearly 100,000 and affect or displace 200 million people](#) annually. As climate change is expected to increase the frequency and severity of disasters in the near future, leveraging [social media](#) data, crowd-sourcing and other means will only become more important.

The potential of data collaboratives

The Malawi partnership is just one manifestation of the concept of [data](#)

collaboratives. We have defined this as a new form of collaboration beyond the public-private partnership model, in which participants from different sectors — including private companies, research institutions, and government agencies — can exchange data to help solve public problems.

While such collaboratives are emerging in a number of sectors and areas, the Malawi case is an example of a particular kind of collaborative. It's what we might call a social media data collaborative.

While much attention has been paid to the impact of social media on politics, much value can be generated from social media data for governing as well, but only when done responsibly.

Users of social media are today disclosing and sharing an unprecedented amount of data. Facebook alone collects 98 unique personal data points from its users, and Twitter processes about 6,000 tweets every second.

With an estimated 2.51 billion social media users across the world, a staggering amount of information is being gleaned about individuals and their interactions from social networking platforms.

There is little doubt that much of the data stored by social media companies could, if made available in a responsible manner, provide groups working for the public interest with new insights and avenues for action. Unfortunately, at present such groups have only limited access to data, and their data science expertise remains similarly limited.

A Missing Maps project event at Facebook's London offices. Credit: OpenStreetMap

Data collaboratives like the Missing Maps project represent a new, contemporary model of corporate social responsibility.

For instance, LinkedIn has established the Economic Graph Research initiative to leverage their data together with a range of third-party researchers to create collective insights for increasing the "economic opportunity for every member of the global workforce." This reflects a growing willingness among companies to provide access to their data to pursue social responsibility goals.

Deploying such models, companies such as Facebook, Twitter and Reddit are no longer simply silent merchants of our personal data. They can use it to serve the public good in a variety of ways. They include:

1) Improved situational awareness and response: In addition to Missing Maps, Facebook has contributed its data to a number of humanitarian projects, with a particular emphasis on improving the accuracy and real-time awareness of humanitarian responses.

The company has shared its commercial building data with the Center for International Earth Science Information Network at Columbia University, for instance. Combined with census data, Facebook's data provides high-resolution information about rural settlements across the globe.

2) Better public service design: Data from social media organisations can help solve everyday problems facing the public.

In its Connected Citizens program, Google Waze shares its crowd-sourced traffic data with ten cities around the world to improve urban transportation. And by accessing municipal inspection data, Yelp displays restaurants' hygiene scores on its business pages, allowing consumers to better judge their quality.

Such data sharing practices between private social media companies and public departments can improve public services and ensure that policies are more responsive to citizens.

3) Enhanced knowledge creation: Social media can be invaluable for researchers looking to access datasets and garner new and innovative insights.

The Digital Ecologies Research Partnership, for instance, allows selected

researchers to extract data from internet communities such as Imgur, Reddit and Stack Exchange to support research on internet social behaviour. And in their Future of Business Survey, the OECD and World Bank use Facebook to deliver surveys and collect data on worldwide business sentiment.

Social media collaboratives can allow scholars to gain access to more granular and up-to-date datasets, generating new research and insights for a variety of applications.

4) Prediction and impact evaluation: Social media data provides valuable information to both anticipate social and environmental problems.

Tweets can be used to predict hurricane damage for instance, or as a tool to evaluate projects after their conclusion.

Facebook partnered with UNICEF to help monitor the reactions and social conversations surrounding its Zika virus public health campaign in Brazil. This allowed the UN body to track the outcome of its initiatives and ensure that its campaign was having the intended effect.

These and other projects suggest that Facebook's trend and status data can provide humanitarian organisations with powerful insights to better coordinate and monitor relief efforts.
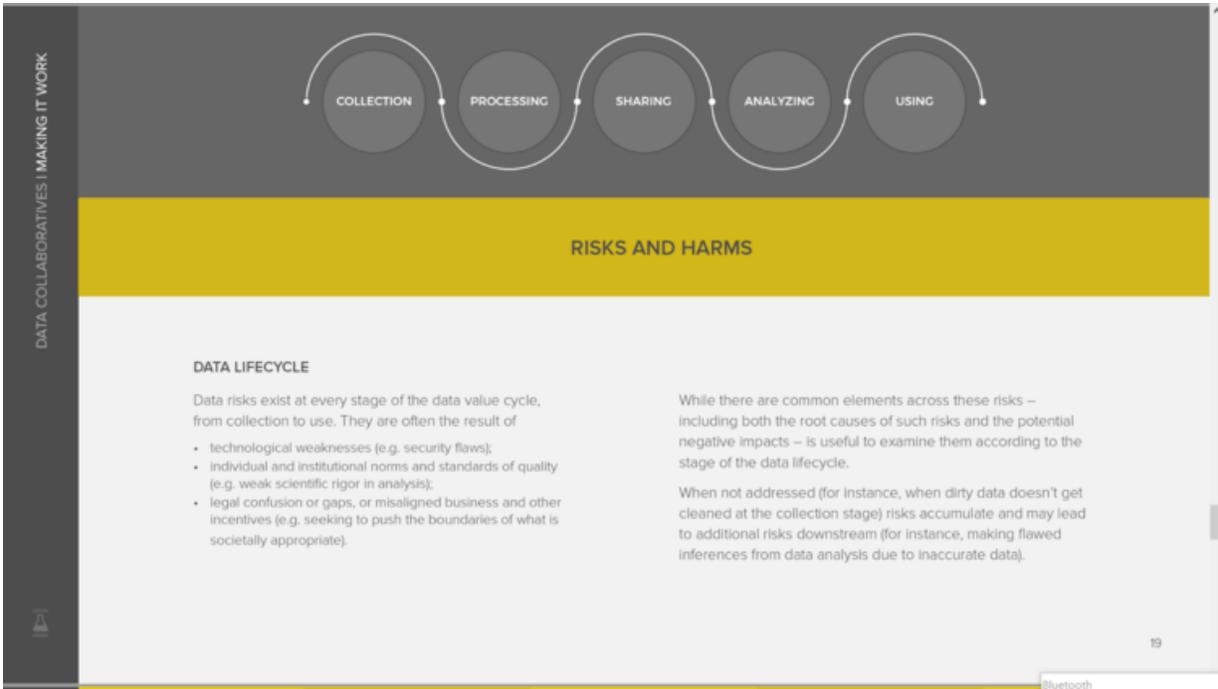
Credit: The GovLab.

Risks of data collaboratives

At any point in the data life cycle, there are inherent risks – from the unauthorised collection of social media information to misrepresenting data through poor analysis and the possible re-identification of individuals once data has been shared.

Such risks are real and ought not to be used as a reason to avoid sharing social media data. Rather, they highlight the need to develop and integrate a data responsibility framework into any data collaborative initiative.

Molly Jackman and Lauri Kanerva from Facebook have argued that when using social media for other purposes:

companies should develop principles and practices around research that are appropriate to the environments in which they operate, taking into account the values set out in law and ethics.

The concept of data responsibility has recently gained traction within a number of industries and sectors, including the social media industry. These latter can create and operationalise responsibility frameworks by employing data stewards – people tasked with determining what and when to share, how to protect, and how to act on available data.

A number of social media organisations have already established separate departments to administer data-sharing projects. Facebook's public policy division, for example, has a review process that focuses on data stewardship.

Other organisations depend on separate, and sometimes independent, intermediaries, such as MIT's Laboratory for Social Machines, which was founded by Twitter's chief media scientist Deb Roy.

Social Machines regularly uses social media data, particularly from Twitter, to support its research and analysis. But, by maintaining its independence and aligning itself with an academic institution, it is able to establish strict guidelines to maintain the ethical rigour of its work.

All of these initiatives are promising, but it is not yet clear that they add up to a comprehensive data responsibility framework or decision tree enabling new ways of working. Such a framework could provide data stewards the means to assess the public value of social media data as well as the risks and harms of sharing it. It could also suggest ways to adequately mitigate this risk.

What's more, it might help achieve the necessary balance between the benefits and risks of sharing, and ensure that the vast amounts of data

being generated by the public every second are ultimately used for the greater good.

More specifically, a generally accepted responsibility framework can help accelerate the emergence of new, innovative data collaboratives, and maximise their potential.

Let's speed up the work initated by bodies such as UN-OCHA, Global Pulse, the International Data Responsibility Group and others, toward building a data responsiblity framework to ensure social media data improves people's lives in a trusted manner.

This article was originally published on The Conversation. Read the original article.

Provided by The Conversation