Beyond DOs and DON'Ts

MAIN CONSIDERATIONS FOR CROWDSOURCING EVIDENCE

Tetyana Bohdanova, Exposing the Invisible In The Summer
About me

- Elections and civil society development specialist, a researcher.
- 10+ years of field-based experience in citizen engagement and electoral transparency work across Eastern Europe and Eurasia
  - including implementation of crowdsourcing projects in restrictive environments

Bohdanova.Tetyana@gmail.com
@TetyUAAna
What is crowdsourcing?

The term crowdsourcing was first coined by Jeff How in a 2006 “Wired” magazine article where he defined it as a new way of sourcing labor enabled by the Internet. Different types of commercial and non-commercial crowdsourcing emerged since. For example:

- Wikipedia is the best example of collective knowledge sourcing
- KICKSTARTER - is an example of crowdfunding projects
- Ushahidi is a popular platform for crowdmapping information
Examples of crowdsourcing

Crowdsourcing may be used by journalists, in a crisis, or for better governance, accountability, and human rights.

Shadow of a Doubt: Crowdsourcing Time Verification of the MH17 Missile Launch Photo

August 7, 2015 - Russia, Ukraine

While much open source investigation is done through crowd sourcing, the administration of the MH17 website forum has taken a unique approach: a single-blind poll to collaboratively determine the timing of a photograph. While this experiment does not have the scientific rigor of a pharmaceutical trial, the methodology of administrator "bootblac" is worthwhile of some examination.

2015, Bellingcat

Bellingcat's flagship investigation into the downing of the Malaysia Airlines 17 (MH17) passenger airplane in Ukraine in 2014 relied heavily on crowdsourcing:
https://www.bellingcat.com/resources/case-studies/2015/08/07/shadow-of-a-doubt/

2019, CORRECTIV, Germany

Open research into who really owns and controls rents in the city's non-transparent real-estate market via own CrowdNewsroom platform.

2019, Humanitarian Open Street Maps Indonesia

Harry Machmud from Humanitarian Open Street Maps (HOT) crowdmapped handwashing stations in Indonesia to prevent the spread of COVID-19.
Allows tapping into a vast pool of data otherwise inaccessible to organizers
1. Allows engaging diverse contributors
2. Can help save time and costs
3. Opens new avenues for collaboration with contributors and/or others working in the same space

1. Carries the danger of manipulation
2. May require a lot of know how and resources
3. Carries the danger of coming up empty-handed
4. May carry potential risks to organizers and contributors
Weighing risks

Given the challenges arising from verifying crowdsourced data and a risk of the data being corrupted by opposing entities, many organizations chose to use crowdsourcing not as a primary but complementary data collection method.

Crowdsourcing often requires a lot of know how and resources, which will be further explored in the next section.

It may be more feasible to employ other methods that could yield similar results, such as open source intelligence that focuses on utilizing free tools and resources.
2. Setting up a crowdsourcing effort

Before you decide to engage in crowdsourcing, it is important to ask a number of questions:

- Why do you want to crowdsource?
- What are the main ethical considerations when crowdsourcing evidence?
- How verifiable is the data collected?
- Who are your contributors?
2. Setting up a crowdsourcing effort

Before you decide to engage in crowdsourcing, it is important to ask a number of questions:

- Who are your beneficiaries?
- What are the risks?
- Who could become your partner or collaborator?
- What happens with data afterwards?
- Should any credit be given and how?
3. Choosing the right approach

Choosing the right approach depends on your goals

A structured call-out: crowdsourcing specific data in a structured way

VS

An unstructured call-out: crowdsourcing data through open unstructured requests

*Typology by the Tow Center for Digital Journalism of Columbia University
Which approach is better for....?

A: Structured call-out

B: Unstructured call-out

C: A mix and/or both

Example 1: Journalists seeking tips for corruption cases to investigate

Example 2: Humanitarian organization mapping wildfires in the region

Example 3: Human rights defenders collecting evidence of police violence during protests
4. Working with target audience(s)

Successfully engaging the community members of which you want to contribute data is half the success of your crowdsourcing effort.

Guiding questions to ask yourself before doing community engagement (from ProPublica):

- Who are the people you want to respond? Why are they the best community to involve?
- What does the community stand to gain from this? What are the reasons someone would participate?
- What are the reasons someone wouldn’t participate? How do you plan to assuage any concerns and hesitations?
- Who’s most affected? What language do they use to describe the problem? Are they angry? Where do they talk about it? How?
- Who are the most influential people in this community? Have you talked to them? What do they think of your idea?
Choosing the right engagement method

Think about how you are going to reach your audience. You may want to push your engagement call via channels and methods that your target contributors, not you, prefer.

- Consider particular social and political conditions in which you operate. If risk is involved, people would be reluctant to do so if they do not believe that tangible change may come out of it.
- Think about how you can get its members interested in and excited about being a part of your crowdsourcing effort.
- Think about how you may demonstrate results and close the feedback loop with your audience even if contributors remain anonymous.
- Be sure to take any privacy and security considerations into account.
You have heard anecdotes about students of public universities being massively pressured to vote by the school administrators.

You would like to collect concrete evidence of such cases, but do not have any presence on the ground.

What would you do?

Option A
Contact student organizations and set up joint hotlines for students to report pressure

Option B
Publish a call in social media groups of particular universities providing an anonymous way to report cases of pressure

Option C
Use local newspapers to disseminate your call
Community outreach methods & channels

De Correspondent

Dutch reporter Jelmer Mommers used the news site to appeal directly to Shell employees for information about the company’s knowledge of climate change. He invited readers to email him and received internal company documents, reports, and more.

https://thecorrespondent.com/4048/dear-shell-employees-lets-talk/446236241-89c4a65a

ProPublica & NPR

US Adriana Gallardo and Renee Montagne disseminated an online questionnaire aimed at women that experienced life-threatening complications in childbirth. Published on Facebook and Twitter, as well as unconventional places like the crowdfunding site GoFundMe, the questionnaire yielded thousands of responses and resulted in a series of publications.

https://gijn.org/2017/08/14/how-they-did-it-propublicas-maternal-mortality-series/

Anyones Child: Families for safer drug control & the University of Bristol

Documentary makers about drug violence in Mexico launched a free phone line publicized through local partners and invited people across the country to call in with their stories. When they did, callers could also listen to the stories of others. This effort resulted in a multimedia documentary project “Anyone’s Child: Mexico”

http://mexico.anyoneschild.org/
5. Choosing tech tools for crowdsourcing

There is an abundance of easy and secure tools developed by and for human rights defenders, journalists, or citizen watchdogs.

Yet, it is important to choose the right tool that would align with the goals and needs of your crowdsourcing effort, not try to design the crowdsourcing operation around it.
Things to consider

Tech environment

1.

- Connectivity
- Low tech vs high tech
- Computer literacy

Privacy & security

2.

- Encrypted vs unencrypted channels
- Security vs usability
- Anonymity vs verification

Resources

3.

- Tech habits are slow to change
- New tools aren't always best
- Skills and know how needed
Examples of tools

There are multiple secure communication tools that are used by journalists and activists. While no system is 100% secure, there are instruments that attempt to create a more secure environment than normal communication channels (such as telephone, social media, email) provide.

No one tool is best for everyone, so it is important to carefully consider individual circumstances of your prospective contributor(s).

- Signal https://signal.org/
- WhatsApp https://www.whatsapp.com/
- Pretty Good Privacy (PGP) email encryption https://www.openpgp.org/software/
- Protonmail https://protonmail.com/
- SecureDrop https://securedrop.org
- Tella https://tella-app.org/
- FrontlineSMS https://www.frontlinesms.com
6. Verification

Verifying crowdsourced data is extremely important. Depending on the type and format of data that you are attempting to collect, think carefully about how much verification you would want and be able to carry out.

1. Unverified data - some data may be impossible to verify as it can be novel and without many corroborating sources. In this case, think of not verification, but vetting.

2. Partially verified data - decide how much verification you deem “enough” for the data to be publicized.

3. Fully verified data - usually, data that has several corroborating sources.

If you are unable to verify the data but still want to publicize it, provide a clear disclaimer marking your data as “unverified”.

Unverified data - some data may be impossible to verify as it can be novel and without many corroborating sources. In this case, think of not verification, but vetting.

Partially verified data - decide how much verification you deem “enough” for the data to be publicized.

Fully verified data - usually, data that has several corroborating sources.

If you are unable to verify the data but still want to publicize it, provide a clear disclaimer marking your data as “unverified”.
Verifying crowdmapped data

Ushahidi users during the 2010 Haiti earthquake identified the following ways of verifying data, including when plotting it on a map:

- location - is the report coming from the right place?
- reputation - is the source trusted by me or by people who themselves are trusted?
- content comparison/aggregation - via clustering or other methods to discover patterns;
- timing - is the report coming at the right time?

Verifying social media reports

There are a variety of techniques for verifying social media reports, multimedia files, etc., that you may use if this is the type of data you collect.

https://datajournalism.com/read/handbook/verification-1
https://datajournalism.com/read/handbook/verification-3

Setting up a verification process is not an easy task and may require creating a complex decision-tree model and a team of people with particular skills.

Example of a journalists’ effort to monitor and verify information during the 2012 Parliamentary elections in Ukraine:
7. Analyzing data & presenting findings

Present crowdsourced data in an honest and truthful manner, but also think of an engaging format.

- How you are going to analyze and present your findings? The format in which you would want to present your findings may also affect the format in which you would to crowdsource data.
- Describe both your methods of data collection and how you arrived at your conclusions (in case of conducting analysis).
- Clearly state whether and to what extent you have been able to verify the data that you crowdsourced.
- Do not forget to give credit.
Amnesty's international real-time open source monitoring of social media during the elections

https://citizenevidence.org/2021/01/21/5-key-takeaways-from-live-monitoring-the-2020-us-election/
Summary

1. Define purpose
2. Consider ethics, legality, and safety
3. Define audience, format, and duration
4. Identify the best method
5. Effectively engage contributors
6. Identify the right tools
7. Set up a data verification process
8. Analyze data
9. Present findings
QUESTIONS?
Thank you!!

Bohdanova.Tetyana@gmail.com
@TetyUAna