

“Exposing the Invisible” Workshops

Crowdsourcing information for investigations

Crowdsourcing: Trainer’s Tip Sheet

Definition

- **Crowdsourcing** is the practice of obtaining information or input into a task or project by enlisting the services of a large number of people, either paid or unpaid, typically via the internet (source: <https://www.lexico.com/definition/crowdsourcing>.)
- Crowdsourcing has been increasingly used by journalists and activists to conduct investigations, collect crisis information, improve accountability, and document human rights abuses.

Key Steps and Tips

Below is a list of key elements and tips that activists and organizations must consider when engaging in crowdsourcing.

1. Define the purpose

- For instance, do you want to tell a story, raise awareness about an issue, collect data to help other activists or organizations in their work, engage citizens around an important process or event, or conduct an investigation?
- The purpose will define other elements of the crowdsourcing effort.

2. Consider ethics and safety aspects

- Before engaging in crowdsourcing, you may want to consider such aspects as:
 - accuracy of the information you are able to collect,
 - privacy and safety of contributors,
 - property / ownership of the data collected,
 - accessibility of your effort.
- Depending on your context, some of these aspects could also have legal implications, so it may be wise to seek legal advice before launching a crowdsourcing project.

3. Define audience, format, and duration

- A) **Who are your main contributors?**
 - Asking specific contributions from trained activists may look different than trying to engage regular social media users.

- You may also want to consider such demographic characteristics as age, gender, or geographic location of your contributors.
- Consider whether you need data from different subsets of the population to paint the full picture.
- **B) What you would like to do with the collected evidence?**
 - For example, do you want to openly share the raw data or write and disseminate a report based on your findings?
 - Is it critical to make the findings publicly available immediately?
 - This may define the format of the data you want to collect.
- **C) How much time do you need?**
 - The duration of a crowdsourcing effort depends on whether you want to organize it around a particular event or a process.

4. Identify the best method

- In an **open callout**, the public is invited to contact organizers with information via various channels (email, telephone, SMS, online surveys, etc.) to contribute feedback, votes, tips, photos, or any other material they wish to submit. This format usually follows the open data collection format.
- In a **specific callout**, organizers target certain groups with a specific request for information. The information is usually provided in a predefined format and captured in a searchable database.
- Sometimes a **mix** of different methods is needed, especially in large collaborative projects or where evidence needs to be cross-referenced across multiple streams of data.

5. Effectively engage target audience(s)

- Successfully **engaging the community members** whom you want to contribute data is half the success of your crowdsourcing effort. They are far more likely to react and help if they feel motivated and approached with relevant information.
- When thinking about community engagement, consider social and political conditions in which you operate, including any risk that your contributors may face.
- Think about how you can get members of the community interested in and excited about being a part of your crowdsourcing effort.
- Sometimes, this may mean that the crowdsourcing should be preceded by some **awareness-raising and trust-building** work and campaigns.

6. Identify the right tools

- It is easy to feel overwhelmed or get overly excited about using a particular technical tool for crowdsourcing.

- It is important to choose the right **tool** that would align with the **goals** and **needs** of your crowdsourcing effort, without trying to design the crowdsourcing operation around the tool.
- When making a choice, consider:
 - A) The **technical environment** (high vs low tech), including tech habits of target audience(s).
 - B) The **security** and **privacy** needs of your contributors.
 - C) The **costs**, **time**, and technical **skills** needed to set up particular tools.

7. Set up a data verification process

- Verifying crowdsourced data is extremely important.
- Depending on the type and format of the data you aim to collect, think carefully about how much verification you may need and are able to carry out.
 - A) Before developing a verification process, decide **how much verification** you deem “enough” for the data to be publicized.
 - B) **Sometimes data cannot be verified, but can be “vetted”**. That is, while it may be difficult to corroborate, if you are unable to refute it the data may still be worthy of being used.
 - C) If you are unable to verify the data but still want to publicize it (i.e., you have vetted it), **provide a clear disclaimer** marking your data as “unverified”.

8. Analyse data and present collected evidence

- Think about how you are going to **analyse** and **present your findings**.
- It is important to present crowdsourced data in an **honest and truthful manner**:
 - A) When publicizing findings, describe both your methods of data collection and how you arrived at your conclusions.
 - B) On your website and any other materials, clearly state whether and to what extent you have been able to verify crowdsourced data.
 - C) Do not forget to give credit where credit is due.
 - D) Findings need to be presented in an interesting and engaging manner. One way to do it, is to think of a “story” you are trying to tell with your data.

Recommended Resources

for further reading when preparing the workshop

NOTE: these resources can also be shared with workshop participants at the end.

Guides:

- “[Guide to Crowdsourcing](#)”, by Mimi Onuoha, Jeanne Pinder, and Jan Schaffer, 20 November 2015, The Tow Center for Digital Journalism (guide archived with Wayback Machine [here](#).)
- [Listening Post Collective](#), a project by Internews that offers resources, tools, peer-to-peer support and a shared learning space for journalists, newsroom leaders and community groups looking to revitalize their local news and information ecosystems through community engagement. It also provides a [Playbook](#) / guide with concrete steps and tips on how to connect with communities (archived with Wayback Machine [here](#).)

Articles:

- “[Crowdsourcing in investigative journalism](#)”, by Nils Mulvad, 4 April 2011, the Global Investigative Journalism Network (GIJN).
- “[How Muckrakers Use Crowdsourcing: Case Studies from ProPublica to The Guardian](#)”, by Toby McIntosh, 27 November 2017, GIJN.
- “[5 Ways to Crowdfund Easily, Legally & with Quality](#)”, by Jeremy Caplan, 9 May 2010, Poynter.org
- “[Should You Crowdfund Your Next Project? Check Out ProPublica’s Checklist](#)”, by Jess Ramirez, 25 June 2019, GIJN.

Cases and Projects developed with the help of crowdsourcing:

- “[A Reporter Crowdfunded ER Bills, and Now Doctors Are Listening](#)”, by Christine Schmidt, 2 September 2019, for GIJN.
- “[Backstory: How the ABC's biggest crowdsourced investigation exposed failings in aged care](#)”, by Jo Puccini, Flip Prior and Fanou Filali for ABC Investigations, 27 April 2018/25 September 2018, for ABC News Australia.
- “[Bribery in India: A website for whistleblowers](#)”, by Mukti Jain Campion, 6 June 2011, for BBC.
- “[CrowdNewsroom: Using Communities to Assemble Non-Public Data Sets](#)”, by Joseph Lichterman, May 13, 2019, for GIJN.
- “[How Crisis Mapping Saved Lives in Haiti](#)”, by Patrick Meier, 2 July 2012, for National Geographic.
- “[Mapping Handwashing Stations in Indonesia to prevent the spread of COVID](#)”, by Cecilia Hinga, 9 April 2021, for Ushahidi.

- “[Mapping the Powerful: Poderopedia Takes Know-How Across Borders](#)”, by Miguel Paz, 16 May 2014, for GIJN.
- “[Monitoring and Verifying During the Ukrainian Parliamentary Election](#)”, by Anahi Ayala Iacucci, *Verification Handbook 1: A Definitive Guide To Verifying Digital Content For Emergency Coverage*: <https://verificationhandbook.com/>, for European Journalism Centre.
- “[Shadow of a Doubt: Crowdsourcing Time Verification of the MH17 Missile Launch Photo](#)”, by Aric Toler, August 7, 2015, for Bellingcat.
- “[STRIKE TRACKER: Decode how US-led bombing destroyed Raqqa](#)”, Syria, April 2019, by Amnesty International.
- “[Using a Mobile Phone Survey to Investigate South Sudan’s Conflict](#)”, by Carolyn Thompson, 11 May 2020, for GIJN (also in [French](#)).
- “[Wem gehört Hamburg?](#)” (“Who Owns Hamburg?”), 23 November 2018, by CORRECTIV (available in German) – a crowdsourced project about real estate ownership in Hamburg, Germany (the project expanded to other cities).

Platforms and Tools used for crowdsourcing data:

- [CrowdNewsroom](#) – crowdsourcing newsroom platform developed by a German investigative reporting outlet CORRECTIV.
- [Fixmystreet.org](#) – open source FixMyStreet Platform allows easily launching a website that helps people to report street problems like potholes and broken street lights. Problem reports are then sent to authorities for fixing.
- [FrontlineSMS](#) – inexpensive software used by a variety of organizations to distribute and collect information via text messaging.
- [Poderopedia](#) (content in Spanish) – a collaborative data journalism platform that helps to understand the relationships between people, companies and organizations across Latin America. The platform also involves citizens who wish to contribute information.
- [Securedrop.org](#) – an open source whistleblower submission system that media organizations and NGOs can install to securely accept documents from anonymous sources.
- [Ushahidi.com](#) – an open source software application, and a non-profit technology company with staff in nine countries whose mission is to help marginalized people raise their voice and those who serve them to listen and respond better.

Organisations known for working with crowdsourced data:

- [Amnesty International](#) – an international non-governmental organization human rights headquartered in the United Kingdom, which runs the “[Amnesty Decoders](#)” platform used by a global network of digital volunteers that help research and expose human rights violations.

- [Bellingcat](#) – an organisation focused on online citizen investigations that often uses crowdsourcing on social media to gather information, document, and verify events as part of its stories and reports.
- [CORRECTIV](#) – an investigative reporting outlet in Germany.
- [ProPublica](#) – an independent American nonprofit newsroom that produces investigative journalism.
- [The Guardian](#) – a British news and media website owned by the Guardian Media Group.