

Data Cleaning Guidelines

Your reports on your map may be public and you may want to share your datasets. The following is a list of clean data guidelines to consider.

Personal Names

Removing names of people is a first priority. This includes personal names, nicknames and potentially pseudonyms. Data lives online potentially forever so protecting your reporters is important.

Phone Numbers

Personal phone numbers for SMS or tracking reports should be removed. These should not be publicly displayed on a website. If you are including SMS reports in your project, remove these before publishing your reports.

Sensitive Information

There is a wide range of sensitive details about a report that you may want to consider removing. For example, for the [Uchaguzi - Kenyan Elections 2013](#) project, we removed tribal names. The reason we did this was to not incite dangerous speech. We worked in partnership with [Ihub Research](#) [considering their Umati](#) (Monitoring Dangerous Speech project).

Geolocation redaction

There will be cases where the exact geolocation of a report needs to be altered. This is for the safety of the topic. For example, if you are mapping an incident in a conflict zone. Consider not having the exact location but instead of general region.

Resources to consider:

Standby Task Force: [Data Protection Standards 2.0](#)

ICT4Peace - [The potential and Challenges of Open Data for Crisis Information Management and Aid Efficiency](#)

Oxfam: [Evolution of Protection of Civilians in Armed Conflict](#)

UNOCHA - [Humanitarianism in a Network Age](#)

ICRC: [Professional Standards for Protection Work Carried out by Humanitarian and Human Rights actors in armed conflict and other situations of violence](#)

- Necessity & capacity
- Data protection laws
- Do no harm
- Bias/non-discrimination (objective information/processing)
- Quality check/reliability

Woodrow Wilson Center: [Privacy and Missing Persons after Natural Disasters](#)