Data Simulation

Simulations are a big part of emergency and humanitarian work. There are exercises and training to help people prepare for logistics. The purpose of this session is to apply this methodology to 'simulate' data workflows for various topics. The inspiration for this session design came from the work done with the IFRC Health group on problem-solving data protection and data sharing issues. We are preparing two workshop examples: one workshop focused on data sharing and one workshop focused on data protection. This session is designed to "learn by doing" methods rather than providing a textbook or slide presentation. It assumes that peer-to-peer learning can provide a unique forum to 'negotiate change'. It provides participants a means to consider how to train and involve people in the conversation to be 'advocates' for data protection and/or data sharing. See the specific scenario examples below for the goals and critical workshop steps.

People:	4 to 12 people
Time:	60 Minutes
Difficulty:	Medium
Materials:	Need large cards/sticky notes by colour for:

- Roles – Green
- Pipeline – Orange
- Actions (sharing/protection) – Yellow
- Roadblocks – Pink
- Other Colours

Painter's tape (to ensure items stick to the wall)
Flipcharts
Markers

PREPARATION:

- Decide on a scenario that best depicts data sharing or data protection issues (pending on the workshop topic)
- Have the scenario printed out on paper or visually on a slide.
    o Alternatively, ask participants to 'create' a real-world example. (The latter takes time and may not provide a shared room experience if 'too niche'.)
- Ask participants to consider all the actions, questions and risks/opportunities to achieve a project that has data protection or data sharing goals.
    o Data Sharing Workshop - the primary goal is to 'simulate' all the barriers, risks and needs to share data.
    o Data Protection Workshop - the primary goal is to 'simulate' all the hurdles, questions, and issues to address to make a project adhere to data protection guidelines.
- Make a diagram on a whiteboard or use a wall (potentially with paper tacked to the wall). The colours are the 'sticky notes'.
- The 'grid' is to have all moving parts, rather than a rigid box table format. The table can be 'ad-hoc', but with clear columns to start. Sticky notes allow the users to move around the parts, add new items, and consider the stakeholders and transactions to make a data flow simulate data sharing or data protection.
- Displayed at the bottom of the chart are roles and tasks that need to be part of the whole cycle.
- Coding the chart:
    o Orange - Data Pipeline items (note that the 'pipeline' includes revisions based on IFRC needs. See School of Data for the original concepts.)

- o Green - Key roles involved in a project or data flow
- o Yellow - Data Sharing questions/ Data Protection concerns/ Key actions/needs
- o Pink - Big risks, needs, gaps
- o Other colours or dots - used to identify priority items for a particular scenario.
- Have a 'parking lot' area to cite outstanding critical questions.
- Ask people to prioritise questions to be addressed.
- In the last 15 minutes, ask people to debrief- what were the common themes, lessons for 'data sharing' or 'data protection'. Ask them how they would address these issues. For example, what do they think the opportunities/barriers are for successful data protection projects and data sharing projects.

An example of what your wall might look like:

Design	PROTO-TYPE	Find	Get	Verify	Clean	analyse	Present	Project Close	Archive
Project lead	Project lead		Trainers			M & E	Comms officers		IT
Legal			Volunteers				Project analysts/ officers		Audit
IT			Local community						
Managers / senior managers									

SCENARIO PLANNING: DATA PROTECTION

Data Protection is part of every humanitarian data-driven project. The goal of this session is to drive up conversations about data protection issues that arise when working with a particular data set and/or project scenario. Time flows very fast in 'exercise mode'. This session design is a minimum one-hour timeslot.

The 'data simulation' focus is to drive conversation around delivering a data-driven project scenario.

Get people talking about real-world data protection issues. The method uses scenarios as examples: either real-world or illustrative. The interactive component provides the means to visualise the steps and actions to 'simulate' decision-making.

It also aims to drive a conversation around the 'implementation steps' and 'requirements' to protect data.

Example Scenario 1: Call Detail Records
During the Ebola response, some telephone companies provided humanitarian data scientists with Call Data Records. This 'metadata' was to be used to help humanitarians obtain population movements. Mobility during an epidemic could potentially cause further spreading of the disease. What are some of the data protection issues that we need to consider.

International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

2

Source - http://cis-india.org/papers/ebola-a-big-data-disaster

About CDRs

SCENARIO PLANNING: DATA SHARING

Sharing data within an organisation or with other humanitarian actors can help reduce duplication and provide insights to support response. There are as many reasons not to share data as there are to share data. With the growth of the Humanitarian Data Exchange and the increasing pressure to share data, how can we work through the various steps and attributes of sharing data? The Innersource methodology provides some tactics to getting to 'open' and 'sharing'.

Time flows very fast in 'exercise mode'. This session design is a minimum one-hour timeslot.

- Get people talking about real-world data sharing issues. The method uses scenarios as examples: either real-world or illustrative. The interactive component provides the means to visualise the steps and actions to 'simulate' decision-making.
- OR/ Drive a conversation around the 'implementation steps' and 'requirements' to share data.

The session should start with the group defining a typical list of data types that might be shared. Also, they should make a list of what kinds of data should not be shared. This provides a way to ensure that people have a shared journey as they walk through the scenarios. Edit the list as the session continues.

Example Scenario 1: Branch Data
You are given a dataset which has a portion of the Red Cross /Red Crescent branches. It was collaboratively created over time and has no 'owner'. The dataset includes country, city, longitude/latitude, admin code level 1, branch name, a point of contact, email address, and phone number. Not all the fields are complete. What type of data would you share or not share? How would you use this dataset? What are the benefits of sharing this data?

SHARING DATA – BASIC CONSIDERATIONS

The following is a basic checklist for sharing data. What other questions would assist the decision-maker

- ☐ Consent
- ☐ Aggregated? Disaggregated?
- ☐ License/ Format
- ☐ Minimization (Only what you need)
- ☐ Owner/ Data Controller
- ☐ Community Feedback Loop: How Data Was Used

Key Questions
- Who needs the data? What is their role? What is the purpose of sharing?
- Who owns or has access to the data? Is it possible to open the data?
- Who can share the data?
- Is there a record of data sharing in the system and/or for the organisation?
- Is there a terms of service agreement with the party that the data was shared with?
- Is there a terms of service and license for the data?
- What capabilities for import, export and exchange of data are required and in which format?

International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

3