Field Guide to Data Sharing

GLOBAL
FOOD SECURITY CLUSTER
Strengthening Humanitarian Response

OCHA

Contents

Introduction

One of the challenges of the humanitarian community is to foster data sharing and collaboration among multiple agencies and organizations, across multiple levels of public, private, and not-for-profit entities. Successful interagency data sharing and collaboration is based on adopting guiding principles, identifying best practices, and recognizing the challenges. It is worth noting that data collected for use within digital systems and also across other systems, should be accompanied by well-defined and responsible data protection practices that are understood by user groups and management teams.

This high level document has been jointly developed by the humanitarian sector to facilitate the sharing of information amongst organizations. This Information Sharing Protocol has been developed following consultation and working experience of participating bodies, chiefly the OCHA Information Management Sub-Group on Data Sharing and the global Food Security Cluster (gFSC) Technology & Innovation Working Group.

Aim of the guide

The aim of this document is to:

- Provide a framework for the establishment and regulation of working practices between participating and Partner Organizations.
- Facilitate sharing of data between the public, private and voluntary sectors so that affected populations receive the services they need.
- Provide increased efficiency and accountability for stakeholders.
- Consider the business processes and controls needed for the sharing of information.

Definition

Information sharing is the disclosure of personal information from one or more organizations to a third party organization or organizations. Information sharing can take the form of reciprocal exchange of data; one or more organizations providing data to a third party or parties; several organizations pooling information and making it available to each other; several organizations pooling information and making it available to a third party or parties; exceptional, one-off disclosures of data in unexpected or emergency situations[1].

[1] NHS New England Information Sharing Policy – personal information, Section 4.1

1. Practical coordination

To determine how information sharing helps parties fulfill their functions and the desired outcomes consider the following questions:

- Why is the data sharing is required, e.g. what is the purpose of this data share?
- What is the desired outcome of the data share?
- Could the desired outcome be achieved without sharing Personally Identifiable Information (PII[2])?

Parties should consider adopting at inter-agency level, the data protection principles as stated in the CaLP Protecting Beneficiary Privacy guide (which reflect the ISO 11 Global Data Protection Principles)

See also the 8 principles of the UK Data Protection Act (1998) and the Policy of the protection of Personal Data of Persons of Concern to UNHCR

1.1. Audience

Who (stakeholders) is the data shared with?

Stakeholders may include: Individuals; NGOs; Government, International Organizations and Affected Population at the field, regional and global/headquarters level.

For an overview of how information flows between stakeholders consult the Humanitarian Decision Makers Taxonomy.

1.2. Type of data that should be shared

The Data Sharing Protocol includes information about specific datasets that helps determine if the dataset should be cleared and shared. Information on the datasets includes risk, aggregation level, sensitivity, guardian of the dataset, barriers to access and actors who the information should be disseminated to.

Access Provisions: The agreement must define who has what rights to access the data at each level of sensitivity/aggregation/normalization, who has what rights to change or modify the data, and what the methods of data access will be.

1.3. Roles and Responsibilities

Individual Responsibilities

[2] **Personally Identifiable Information (PII):** Any data that could potentially identify a specific individual. Any **information** that can be used to distinguish one person from another and can be used for de-anonymizing anonymous data can be considered **PII**.

Every individual working for the organizations listed in this guide is personally responsible for the safekeeping of any information they obtain, handle, use and disclose.

Organizational Responsibilities

Each Partner Organization is responsible for ensuring that their organizational and security measures protect the lawful use of information shared under this Protocol. Organizations are expected to promote staff awareness of the major requirements of Information Sharing. This will be supported by the production of appropriate guidelines where required that will be made available to all staff via the Intranet sites and/or via other communication media.

1.4. Platforms for data sharing

Data sharing is possible through the use of data sharing platforms. Selecting the appropriate data sharing platform may depend on if the data sharing is a one way sharing or a two way sharing.

New data sharing platforms need to have[3]:

a) Flexibility – easy and quick adaptation to new and different situations
b) Scalability – adoption at scale
c) Extensibility – incorporation of new innovations without substantial impact on current systems or workflow.

Examples of Data Sharing Platforms:

Humanitarian Data Exchange: The Humanitarian Data Exchange (HDX) is an open platform for sharing data. The goal of HDX is to make humanitarian data easy to find and use for analysis. Launched in July 2014, HDX has been accessed by users in over 200 countries and territories. Watch the launch animation or introductory screencast to get started.

Geonode: GeoNode is a web-based application and platform for developing geospatial information systems (GIS) and for deploying spatial data infrastructures (SDI). It is designed to be extended and modified, and can be integrated into existing platforms.

1.5. Coordination table

Coordination between field and remote teams:

The recent experience from the Ebola response provides a useful starting point for data coordination between field and remote teams. Its roots are in the Coordinated Data Scramble Community of Interest established in 2011 during the OCHA Wash-up meeting after Libya and Japan responses. During the Ebola

[3] CTP Interoperability Challenges and State of Play – Thoughtworks 2015

response the DHN established an IM/GIS Skype group for facilitating communication on data issues, and a coordination spreadsheet (see DHNetwork example). This procedure was discussed at the recent DHN Summit in November 2014, and agreement was reached that formalization of the Coordinated Data Scramble was a high priority. As a result a Concept Note was developed. The Data Release Protocol presented by REACH provides some excellent ideas for how data sensitivity (a common reason for limiting data sharing) could be incorporated into the coordination spreadsheet template.

1.6 Data and Information Storage

Under Principle 6 (Security) of the CaLP Protecting Beneficiary Privacy guidelines, parties need to:

- "Ensure that organizational and programme systems are in place to ensure beneficiary data is securely stored, e.g. where possible programme staff should liaise with in-house IT staff on information security"
- "Ensure that digital storage systems are encrypted and password protected, and if hard copies of records are retained that include beneficiary data, make sure these records are kept in a secure place"

2. Technical standards

2.1 Data collection

In order to be appropriate for sharing and to be aggregated data collection should meet minimum requirements including:

Sex and Age Disaggregated Data

Humanitarian responses that are inclusive and accountable to affected populations acknowledge differences linked to gender, age, and diversity, including disability and other vulnerabilities, and are informed by the analysis, at a minimum, of sex and age disaggregated data (SADD). **Projects that analyse and take into consideration the needs, priorities, capacities, constraints and risks of both the female and male population of all ages, from all diversity backgrounds, are far more likely to improve the lives of affected populations.** Not analyzing and addressing gender and age needs and capacities puts at stake the efficiency and effectiveness of the response as it will not adequately address the needs of a large part of the affected population and tends to ignore important power dynamics, and could unwittingly cause harm to those we aim to assist.

In an emergency, the initial focus is on primary needs and on meeting these through the delivery of aid, such as emergency food security assistance, as quickly as possible[4]. In the aftermath of an emergency, affected communities will need to restart agricultural and other livelihoods activities as soon as possible[5]. However, distributing food assistance (i.e. either food rations or cash/vouchers) – directly or through

[4] https://www.humanitarianresponse.info/en/operations/afghanistan/document/food-security-1-food-assistance-gender-marker-tip-sheet

[5] https://www.humanitarianresponse.info/en/operations/afghanistan/document/food-security-2-agriculture-and-livelihoods-tip-sheet-september-2012

food/cash-for-work or food/cash-for-training projects – will not automatically guarantee their optimal use or a positive impact on individuals or on the affected population; only a gender and age sensitive, participatory approach at all stages of the humanitarian programme cycle can help ensure that an adequate and efficient response is provided[6]. **In order for a good assistance project to have a positive impact, women, girls, boys and men must be equally and meaningfully involved in the process**.[7]

A needs assessment (e.g. MIRA/ EFSA/ CFSAM, etc.) is the essential first step in providing emergency food assistance and in planning agriculture/livelihoods programming that is effective, safe and restores dignity. A gender and age sensitive analysis (based on primary and secondary sex and age disaggregated data), that also takes into account the socio-cultural context of the emergency, is necessary to understand the social and gender dynamics that could help or hinder the effectiveness of the response. The gender and age analysis during the needs assessment will identify gender gaps, such as unequal access to food assistance or agriculture / livelihoods services for women/girls and men/boys that need to be addressed. The gender and age analysis should then inform the relevant sections of the Humanitarian Needs Overview (HNO) and the Humanitarian Response Plan (HRP), as well as the activities of selected projects. As well, the project's outcomes should capture the change that is expected for different female and male beneficiary groups (e.g. young boy, adolescent girl, adult man, older woman, etc.) and be reflected in the monitoring framework.[8]

2.2 Aggregating data

When dealing with aggregated data the following conditions should be considered:

- Where aggregated data is stored. This includes the physical and logical location of the data, and where copies of the data are stored.
- Who can access aggregated data
- Sunset rules and handing of history. Unless there is compelling and clearly documented justification, organizations (and their third party partners) should not keep beneficiary data longer than necessary for the purposes for which it was collected. There needs to be clear processes to ensure the secure disposal and destruction of data by all parties who have accessed it.[9]
- Handling of data conflicts[10]. A clearly documented and defined process including arbitration in case of conflict between 2 or more parties.

[6] For information: On average, men comprise 57% and women 43% of the agricultural labor in developing countries6. According to FAO data, if women had the same access to inputs as men, agricultural production worldwide would increase by 2.5 to 4 percent and the number of people suffering chronic hunger would decline by 12-17 percent. Men and women work as partners in most subsistence and small-holder farming, sharing some tasks but often performing activities that the other sex does not.

[7] Food Security 1 (Food Assistance) & 2 (Agriculture and Livelihoods) Gender Marker Tip Sheets

[8] Key Messages: Gender & Age Sensitive Humanitarian Response Contributes to AAP, gFSC

[9] Adapted from "Protecting Beneficiary Privacy" – CaLP

[10] CTP Interoperability Challenges and State of Play – Thoughtworks 2015, "Strategies for Semantic Interoperability, pgs. 36-37.

2.3 Digital Identity

Humanitarian agencies and private sector players such as those in the payment card industry need to collaborate to develop the framework for the development standard identity for the issuance of a portable identity to beneficiaries by aid organizations. This digital identity (DIGNITY) standard will:

1. Ultimately provide beneficiaries with a portable digital identity that they can use to identify themselves without connectivity and expensive hardware. The digital identity would be accepted by humanitarian players as well as third party service providers (such as Financial Services Providers) who may be involved in the service delivery of aid to beneficiaries.

2. Provide transparency on and accountability for the reliability of the identity document issued for the client.

3. Limit the collect of information to the minimum essential dataset needed to be able to provide assistance to beneficiaries.

This standard will help determine

1. What information to collect and how to collect the same.

2. How to assess and communicate the reliability of the identity information.

3. Management of the collected identity information.

4. The life span of the identity issued to clients.

5. Help inform beneficiary information interoperability rules and standards.

2.4 Standards datasets

OCHA maintains Common Operational and Foundational Operational Datasets, common datasets needed for response in humanitarian emergencies, as well as the governance model for the management of the data (i.e. accountabilities & responsibilities).

- **Common Operational Datasets (CODs)** are critical datasets that are used to support the work of humanitarian actors across multiple sectors. They are considered a de facto standard for the humanitarian community and should represent the best-available datasets for each theme.
- **Fundamental Operational Datasets (FODs)** are datasets that are relevant to a humanitarian operation, but are more specific to a particular sector or otherwise do not fit into one of the seven COD themes.

2.5 Data formats

The schemas for these data sets are generally not standardized across different actors nor are the mechanisms for sharing the data. In the best case, this results in a **significant delay between the collection of data and the formulation of that data into a common operational picture**. In the worst case, information is simply not shared at all, leaving gaps in the understanding of the field situation.

Most prior attempts to address this problem have focused on building new tools in the form of databases or forms for collecting this information from humanitarian actors in a standardized way. These attempts have had limited success because they require humanitarian organizations to:

- change their internal information management processes or
- add additional work to already over-burdened staff who must fill out the standard form (online or otherwise)

OCHA, in collaboration with the Preparedness and Prioritization Community of Interest is undertaking an initiative to build a data exchange language to address this problem in a new way based on an approach that has been successfully used in other domains. Key to the success of this approach is that it **does not require changes to existing information management tools and procedures** in use in a given humanitarian organization. Instead, an open export format called the Humanitarian Exchange Language is defined that allows organizations to publish their data in a machine-readable format.

The Humanitarian Exchange Language (HXL) is a cooperative data standard that has a special focus on hashtags for reporting 3W (responder's activities) and humanitarian profile (the needs of affected people) data. The HXL is currently as available as the HXL standard version 1.0 beta. For more information see project overview.

2.6 Standards and interoperability

Semantics- Three types of semantic meaning need to be aligned when exchanging data:

- Meaning of Terms: What does this variable mean
- Meaning of Values: What does the variable measure
- Hierarchy and Relationships: Data gains value by being linked to other terms.

2.7 Confidentiality and Disclaimers

There should be a disclaimer covering the accuracy of the data. For example on maps produced by UN agencies the following Disclaimer should be clearly readable, "The boundaries and names shown and the designations used on this map do not imply official endorsement or acceptance by the United Nations."

Description of the data along with appropriate metadata (data that describes and gives information about other data) should also be provided. Metadata is vital is ensure that the copyright of the data are preserved and any specifics on the use of the data are maintained. To balance the need for metadata against the burden of creating and maintaining it, OCHA recommends a 3-tiered system of metadata.

3. Ethics: protecting privacy & minimizing risk

3.1. Minimum standards to apply where a relevant legislation is missing

Partner agencies should, as a rule, conduct Privacy Impact Assessments (PIA) to avoid duplication of efforts and to ensure common understanding of risks and risk mitigation strategies. PIA's become more and important in scenarios where relevant legislation is missing as it helps minimize the risk to all stakeholders including beneficiaries while instilling public confidence in the organisation and reducing reputational risk. The use of PIA's is in line with the Principle of "Protection by Design"[11].

More information on PIA's can be found on pages 28 -29 of the Policy on the Protection of Personal Data of Persons of concern to UNHCR.

From a practical perspective, it is also very important to set some operative standards, in particular when relevant legislation is missing (as it often happens). This will provide support to several NGOs without a legal office or any idea about how to formally deal with these topics.

3.2. Consent protocols

Develop (or use a pre-existing) standard form (consent) with authorization for the use of data collected. It should be a template that can adapted and translated into different languages as needed. The form should have provisions for engaging with the beneficiaries and communities as well as clearly stipulate how the collected data will be used, who the data will be shared with and for how long the data will be retained.

To formulate a consent protocol the following information should be considered:

- What should never be shared? E.g. individual registration data
- What will be dependent on the situation? According to the situation a consent protocol may result in the following outcomes:
    - Never to share data
    - Data to share only under the following conditions…
    - Data to share always apart the following situations…
    - Data free to share

[11] CaLP – Protecting Beneficiary Privacy

- Informed consent: is it really free and informed?

Part A of Annex 2 of the *Protecting Beneficiary Privacy : Principles and operational standards for the secure use of personal data in cash and e-transfer programmes"* guideline document has model clauses on Beneficiary Notice and Consent. The Annex has which has clauses which will be shared with the beneficiary in order to obtain their consent to collect their data. The clauses explain:-

- Why the data is being collected and what it will be used for.
- How long the data will be kept for
- The type of data that will be collected
- Whom the data will be shared with and the circumstances thereof
- Explanation of how the agency will protect the data
- Rights of the beneficiary with regards their data

3.3 Restrictions

All shared information, personal or otherwise, must only be used for the purposes specified at the time of disclosure as defined in the relevant Information Sharing Agreement

3.4 Anonymized Data

Anonymization is the process of turning data into a form which does not identify individuals and where identification is not likely to take place. This allows for a much wider use of the information. For example to ensure confidentiality and protect organizations distributing humanitarian assistance in sensitive geographical areas, instead of listing organization names use a coding system to identify the organization See the following link for more information: https://ico.org.uk/for-organisations/guide-to-data-protection/anonymisation/

4. Legal issues

4.1. Licensing types

- Data is collected in a fair and lawful manner (It is a tricky aspect: if data is collected in a country where there is no legislation regarding data collection, and then are processed or used in a country where there is (and data collection did not fulfill its requirements), what happens?)
- Data is collected for a specific and legitimate purpose. Only collecting the minimum set of data needed for humanitarian agencies to do their work.
- Data access and liability in case of data breach
- Data ownership in environment where data is shared.
- Assurance from third parties on safeguarding of transferred data
- Full accountability for access to and use of data.

- Rights of beneficiaries (e.g. right to information access, right to correction of incorrectly captured data, right to claim compensation for any breaches of policy
- Principle 6 - Data Protection Principles – Data protection Act (UK)

5. Generic security

Measures to be considered include maintaining physical security by having servers and computer hardware and peripherals in a secure location with very strict access control. All technology assets should be tagged as part of the organizations asset management practices. In addition industry best practices, such as the use of firewalls, Secure Sockets Layer (SSL) and Virtual Private Networks (VPN) should be implemented as part the overall security plan.

5.1. Secure Electronic Devices

Computers

First and foremost, it is important to reduce your computer`s vulnerability to hackers and malicious software, otherwise known as malware, such as viruses, trojans, and spyware. Otherwise, any attempt to try and guarantee the security of your communication and computer files will not be successful.

How-to Booklet: This booklet explains how to maintain your software and use tools such as Avast, Spybot and Comodo Firewall to protect your computer against malware infections and hacker attacks.

Mobile Phones

Mobile phones are an integral part to how we access, store and share information. The Information stored on and sent by mobile phones is insecure. "The way the mobile networks operate, and their infrastructure, are fundamentally different from how the internet works. This creates additional security challenges, and risks for users' privacy and the integrity of their information and communications." Advances in technology now mean that mobile phones can provide services and features similar to desktop or laptop computers. These smartphones offer many new ways to communicate and capture and disseminate along, resulting in additional security challenges.

How-to Booklet 1: This booklet explores the security challenges associated with using basic mobile phones and what a user can do in light of these issues in order to use mobile phones as securely as possible.

How-to Booklet 2: In this booklet you will explore the additional security challenges posed by using smartphones. This booklet will teach you some basic setup procedures for securing information and communication your smartphone. You will learn specific precautions related to common uses of smartphones such as emailing, capturing media, accessing the internet and storing information.

Security for other electronic devices including radios, satellites and biometrics

Humanitarian actors must be aware of local factors that may have an impact on the security risks in relation to the field sharing of data. Some of these include:

- Surveillance of mobile communications and government control of phone and internet connections.
- Vulnerabilities of technologies (such as mobile phones, radios and satellites) used in the collection, storage and transfer of data.

In addition, parties need to proceed with caution in relation to the collection, use and sharing of biometric data as it can have serious implications for beneficiary data privacy. The collection of biometric data should not by default. Parties should consider alternatives first and there needs to be a full and comprehensive analysis to justify the use of biometrics and how the privacy issues associated with biometrics are outweighed by using biometrics. Once a determination has been made that biometric data is to be collected, that parties have the capacity to treat such data with extreme care and implement procedural and technical safeguards which include

- Direct capture, encryption and independent storage.
- Biometric data only use to confirm eligibility for benefits
- Access to biometric data must be strictly controlled [12]

5.2. Secure communication

Secure Passwords

Passwords are often the first, and sometimes the only, barrier between information and anyone or anything that might want to read, modify, or destroy it without permission.

How-to Booklet: this booklet will teach you the elements of secure passwords, tricks for remembering complicated passwords, and how to use tools such as KeePass to store your numerous passwords instead of having to remember them all.

Securing your Email

If an email message is intercepted on the way to a recipient, the contents of the message can be read. Your *Internet Service Provider (ISP)* is the first recipient of an email message as it begins its journey to the recipient. Similarly, the recipient's *ISP* is the last stop before your message is delivered. And, because the Internet is just one large, worldwide network that relies on intermediary computers to direct traffic, many different people may have the opportunity to intercept a message in this way. Unless you take certain

[12] Privacy and Big Data Institute. “Humanitarian Cash Transfer Programs and Beneficiaries:
Know Your Customer Standards and Privacy Recommendations”

precautions, your messages can be read or tampered with at either of these points, or anywhere in between.

How-to Booklet: In this booklet you will learn important steps that you can take in order to increase the security of your email communication.

Secure Instant Messaging

Instant messaging and voice communication is not normally secure. However, there are programs that can help secure the privacy of chat sessions but it is necessary that all of your instant messaging contacts take the same security precautions.

Skype is one of the most common instant messaging and voice communication tool that supports calls to landlines and mobile phones. According to Skype, it encrypts both messages and voice calls, however it is important to note that this only happens when both communicating sides are using Skype programs. Skype does not encrypt calls to phone or text sent as SMS messages. As Skype is a closed-source program, it is impossible to do an independent audit and evaluation of its proclamations about encryption, it is thus impossible to know how well Skype is protecting the users and their communication. If are using or plan to use Skype there are important precautions you can take to help improve the security of your communication through skype which are listed here.

Google Talk "Hangouts": is another common instant messaging and voice communication tool. It ises an open protocol (XMPP) with SSL for setting security levels. It is important to familiarize yourself with the various security and privacy settings to ensure you have taken the necessary steps to increase the security of your communication.

- Protecting myself while on Google Talk
- Chatting off the record

Pidgin is a chat program which lets you log in to accounts on multiple chat networks simultaneously. This means that you can be chatting with friends on MSN, talking to a friend on Google Talk, and sitting in a Yahoo chat room all at the same time. Pidgin runs on Windows, Linux, and other UNIX operating systems. Pidgin is compatible with the following chat networks out of the box: AIM, ICQ, Google Talk, Jabber/XMPP, MSN Messenger, Yahoo!, Bonjour, Gadu-Gadu, IRC, Novell GroupWise Messenger, Lotus Sametime, SILC, SIMPLE, MXit, and Zephyr. It can support many more with plugins. Pidgin is multiprotocol meaning it can use both open (e.g. XMPP) and private protocols (e.g. MSN).

5.3 Apps and add-ins that can help improve security

Security in-a-Box is a guide to digital security for activists and human rights defenders throughout the world. The Tactics Guides cover the basic principles of digital security and recommended tools, including advice on how to use social networking platforms and mobile phones more safely. The Tool Guides offer step-by-step instructions to help you install and use the most essential digital security software and services.